

STIC-Biotech/ChemLib

86811

From: Yu, Misook
Sent: Friday, February 14, 2003 9:15 AM
To: STIC-Biotech/ChemLib
Subj ct: 09865879

RECEIVED

FEB 14 2003

CON/CHEM. EN. ISIL.
(STIC)

please search SEQ ID NO:1

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/20
Date Completed: 2/29
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Yu, Misook

To: STIC-Biotech/ChemLib
Subject: 09865879

please search SEQ ID NO:1

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

GenCore version 5.1.4-P5-A578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 08:00:04 ; Search time 4712 Seconds

(without alignments)
1154.429 Million cell updates/sec

Title: US-09-865-879-1

Perfect score: 1806

Sequence: 1 999gattcgttcttctcct.....9gcgcgccccggccgccag 1806

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcg_hum.*
40: em_hcg_mus.*
41: em_hcg_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806	100.0	1806	AX323409	AX323409 Sequence
2	1795	99.4	10884	AX409699	AX409699 Sequence
3	1795	99.4	10884	HUM1BP3	M35878 Homo sapien
4	1685.4	93.3	69887	AC091524	AC091524 Homo sapi
5	1141.4	63.2	8700	AX345904	AX345904 Sequence
6	1021.2	56.5	8700	AX345905	AX345905 Sequence
7	190	10.5	13720	AF305712	AF305712 Bos tauru
8	82.6	4.6	107161	AC109771	AC109771 Rattus no
9	78.6	4.4	7218	I66494	I66494 Sequence 14
10	74.4	4.1	207420	AC078884	AC078884 Mus muscu
11	74.2	4.1	68868	AC101770	AC101770 Mus muscu
12	70	3.9	280595	AC123860	AC123860 Mus muscu
13	68.2	3.8	187521	AC125645	AC125645 Rattus no
14	68.2	3.8	213985	AL591129	AL591129 Mouse DNA
15	68.2	3.8	214511	AL806529	AL806529 Mus muscu
16	67.8	3.8	112283	AL732574	AL732574 Mouse DNA
17	67.8	3.8	174099	AC109560	AC109560 Rattus no
18	67.6	3.7	152996	AC008177	AC008177 Homo sapl
19	67.6	3.7	176108	AC125139	AC125139 Mus muscu
20	67.6	3.7	201893	AL833778	AL833778 Mus muscu
21	67.6	3.7	209992	AC117827	AC117827 Mus muscu
22	67.4	3.7	177145	AC114665	AC114665 Mus muscu
23	67.2	3.7	204431	AC096927	AC096927 Rattus no
24	67	3.7	72358	AC101632	AC101632 Mus muscu
25	67	3.7	149211	AC102624	AC102624 Mus muscu
26	67	3.7	200885	AC119856	AC119856 Mus muscu
27	66.8	3.7	146091	AC099260	AC099260 Rattus no
28	66.8	3.7	162466	AC119298	AC119298 Rattus no
29	66.6	3.7	168073	AC129201	AC129201 Mus muscu
30	66.4	3.7	972	PM12A12G	PM12A12G
31	66.4	3.7	150305	AC114705	AC114705 Rattus no
32	66.4	3.7	176107	AP002393	AP002393 Homo sapl
33	66.2	3.7	100740	AC131216	AC131216 Rattus no
34	66.2	3.7	193273	AC104324	AC104324 Mus muscu
35	66.2	3.7	201650	AC119304	AC119304 Rattus no
36	66.2	3.7	203343	AC131133	AC131133 Rattus no
37	66.2	3.7	242694	AC115360	AC115360 Mus muscu
38	66	3.7	147405	AC118576	AC118576 Didelphis
39	66	3.7	168002	AC115476	AC115476 Rattus no
40	66	3.7	181492	AC099587	AC099587 Mus muscu
41	66	3.7	201838	AC124508	AC124508 Mus muscu
42	65.8	3.6	248201	AC115293	AC115293 Mus muscu
43	65.8	3.6	281892	AC125486	AC125486 Mus muscu
44	65.6	3.6	88064	AC022707	AC022707 Homo sapl
45	65.6	3.6	169709	AC105300	AC105300 Mus muscu

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AX323409	AX323409	Sequence 1 from Patent WO0192578.	AX323409	AX323409.1	GI:18094172	human.	Homo sapiens	1	Roninson, I.B., Dokmanovic, M. and Chang, B.D.	Reagents and methods for identifying and modulating expression of genes regulated by retinoids
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			

JOURNAL Patent: WO 0192578-A 1 06-DEC-2001;
Board of Trustees of the University of Illinois (US)
Location/Qualifiers
FEATURES 1. 1806
Source /organism="Homo sapiens"
/db_xref="taxon:9606"
/note="IGFBP-3 NCB1 acc. number : M35878.1"

BASE COUNT 467 a 437 c 517 g 385 t
ORIGIN

Query Match 100.0%; Score 1806; DB 6; Length 1806;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGATTCGTTTGGTTCCTTCATTTTCCATGAATCAGAGTCTGTTGGGTG 60
DB 1 GGGGATTCGTTTGGTTCCTTCATTTTCCATGAATCAGAGTCTGTTGGGTG 60
QY 61 CAACGCACTACTAGAGAGGAGTGCATCAAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 CAACGCACTACTAGAGAGGAGTGCATCAAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 ACGGTTTCCTGTAAACAGGTTGAGTGTAGCCACAGGCTGAGCAGTGGAGAGAG 180
DB 121 ACGGTTTCCTGTAAACAGGTTGAGTGTAGCCACAGGCTGAGCAGTGGAGAGAG 180
QY 181 CATAGAGAGAGAGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CATAGAGAGAGAGAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TGGTCTCTTTTATTTTGTAGAGTCCATTTATTTATTTATTTATTTATTTATTT 300
DB 241 TGGTCTCTTTTATTTTGTAGAGTCCATTTATTTATTTATTTATTTATTTATTT 300
QY 301 ATTGATCTTTCCATTTCTAAAGAGAGATATATATATTAAGAGAGCTATATGAT 360
DB 301 ATTGATCTTTCCATTTCTAAAGAGAGATATATATATTAAGAGAGCTATATGAT 360
QY 361 GGGGAGAGAGTGTGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GGGGAGAGAGTGTGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GACAGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GACAGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TAACTTCAACCAAGGCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 TAACTTCAACCAAGGCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 TACTCCAAACTGCAGAGAGGCTCTTCAAGTCACTATCCACTCTTAAGGCAAG 600
DB 541 TACTCCAAACTGCAGAGAGGCTCTTCAAGTCACTATCCACTCTTAAGGCAAG 600
QY 601 ATTTCAGAGGCTTCAAGAGAGTCTCTCCCGGAGAGTCAACGCTTCCACTCACC 660
DB 601 ATTTCAGAGGCTTCAAGAGAGTCTCTCCCGGAGAGTCTCAACGCTTCCACTCACC 660
QY 661 CCCACAAACTCTTTGGAGAGAGTGCCTTGAAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 CCCACAAACTCTTTGGAGAGAGTGCCTTGAAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CCAGGCTCTGTGTGTGTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 721 CCAGGCTCTGTGTGTGTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 GTGCCAGAGAGAGAGAGAGAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 GTGCCAGAGAGAGAGAGAGAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 TACTGAG 900
DB 841 TACTGAG 900

QY 901 AATAGAGAGTGCCTTCCAAAAATGACAAATGAAGAGAGAGAGAGAGAGAGAG 960
DB 901 AATAGAGAGTGCCTTCCAAAAATGACAAATGAAGAGAGAGAGAGAGAGAGAG 960
QY 961 AATTACGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 AATTACGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 CTCATTTGAAG 1080
DB 1021 CTCATTTGAAG 1080
QY 1081 AATATTTGAAG 1140
DB 1081 AATATTTGAAG 1140
QY 1141 CTTCGAGTGTAG 1200
DB 1141 CTTCGAGTGTAG 1200
QY 1201 AGGCCCCGGGAG 1260
DB 1201 AGGCCCCGGGAG 1260
QY 1261 GGGGTAAATTTAAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 GGGGTAAATTTAAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GTCCCAAGAGTGTGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 GTCCCAAGAGTGTGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 AATGCGAG 1440
DB 1381 AATGCGAG 1440
QY 1441 CGAG 1500
DB 1441 CGAG 1500
QY 1501 GAGTGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 1501 GAGTGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 CTTTGGTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 CTTTGGTCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 CCGGCGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1621 CCGGCGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 GGGGCGGAG 1740
DB 1681 GGGGCGGAG 1740
QY 1741 CGTTCCGAG 1800
DB 1741 CGTTCCGAG 1800
QY 1801 GCCCAG 1806
DB 1801 GCCCAG 1806

RESULT 2
AX409699 10884 bp DNA linear PAT 14-JUN-2002
LOCUS AX409699
DEFINITION Sequence 2346 from Patent WO0229103.
ACCESSION AX409699
VERSION AX409699.1 GI:21442404
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Alvaraz, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2346 11-Apr-2002;
GENE LOGIC INC (US)
FEATURES
source 1..10884
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M35878"
BASE COUNT 2796 a 2578 c 2737 g 2773 t
ORIGIN
Query Match 99.4%; Score 1795; DB 6; Length 10884;
Best Local Similarity 99.9%; Fred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGATGCTGTTGTTGCTTCATTTTCCATGAAATCAGAGATCTGTGCGTGT 60
DB 101 GGGATGCTGTTGTTGCTTCATTTTCCATGAAATCAGAGATCTGTGCGTGT 160
QY 61 CAACGACATCTAGAGAGAGGTGATACAGAGAAAG- AACAGCAGCAGATTATG 119
DB 161 CAACGACATCTAGAGAGAGGTGATACAGAGAAAG- AACAGCAGCAGATTATG 220
QY 120 CACGGTTCCGTAAACAAGTTGAGTGTAGCCACAGCCTGAGCAGTGGGAGAGAG 179
DB 221 CACGGTTCCGTAAACAAGTTGAGTGTAGCCACAGCCTGAGCAGTGGGAGAGAG 280
QY 180 TCATTAAGAAATGACGCTGCTGGGCTTCGTCACCCCGGCGCCATGTTGTTGCT 239
DB 281 TCATTAAGAAATGACGCTGCTGGGCTTCGTCACCCCGGCGCCATGTTGTTGCT 340
QY 240 TTGCTCTCTTTTATTTGTAGAGTCCATTTATTTATTTATTTAGTACAGAGGAGCA 299
DB 341 TTGCTCTCTTTTATTTGTAGAGTCCATTTATTTATTTATTTATTTAGTACAGAGGAGCA 400
QY 300 AATTGATCTTTCCATTTTAAAGAGAGATATTTATTTATTTAAAGAGAGCTGATATGAT 359
DB 401 AATTGATCTTTCCATTTTAAAGAGAGATATTTATTTATTTAAAGAGAGCTGATATGAT 460
QY 360 GGGGAGAGAGTGTGACAGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
DB 461 GGGGAGAGAGTGTGACAGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520
QY 420 GGACAGAGAGAGACAGTGGGAGAGATGATTAGAGAGAGAGAGAGAGAGAGAGAG 479
DB 521 GGACAGAGAGAGACAGTGGGAGAGATGATTAGAGAGAGAGAGAGAGAGAGAGAG 580
QY 480 TTAATTCACCAAGGCTGCTGCTGAGAGGGAGTGGAGAGAGAGAGAGAGAGAGAG 539
DB 581 TTAATTCACCAAGGCTGCTGCTGAGAGGGAGTGGAGAGAGAGAGAGAGAGAGAG 640
QY 540 CTACTCCAAACTGCAAAAGGCTCTTCAAGTCACTATTCACCTCTTAAGGCAAGCCTGC 599
DB 641 CTACTCCAAACTGCAAAAGGCTCTTCAAGTCACTATTCACCTCTTAAGGCAAGCCTGC 700
QY 600 AATTTCACAGAGTGTGAGAGAGTCTCTCCGCGGAGTCTCACCGCTTCCACTCCAG 659
DB 701 AATTTCACAGAGTGTGAGAGAGTCTCTCCGCGGAGTCTCACCGCTTCCACTCCAG 760
QY 660 CCCCAACAAGCTTTGAAAAAGTGCCTTGAATAATTTAATCCTCAATCCGAGAGC 719
DB 761 CCCCAACAAGCTTTGAAAAAGTGCCTTGAATAATTTAATCCTCAATCCGAGAGC 820
QY 720 ACCAGAGTCTCTGTTGGTCAACGAGAGAGGGGTGCGCAGACAAAGCTGAAGAACTCG 779
DB 821 ACCAGAGTCTCTGTTGGTCAACGAGAGAGGGGTGCGCAGACAAAGCTGAAGAACTCG 880

QY 780 AGTCCAGAGAAAGCCGACAGAGATTACAGCAGCTCAGCGGCAATTCGCCCCGAACT 839
DB 881 AGTCCAGAGAAAGCCGACAGAGATTACAGCAGCTCAGCGGCAATTCGCCCCGAACT 940
QY 840 TTACTGAAAGGCTTTAGATTGAGAGATTAAGCTAATTCACAGCAGCTGAGAAATAC 899
DB 941 TTACTGAAAGGCTTTAGATTGAGAGATTAAGCTAATTCACAGCAGCTGAGAAATAC 1000
QY 900 TTAATGCAAGTGGCTTCAAAAAATGACAAATGAAATTCCTTTAAGAGACTTTGGT 959
DB 1001 TTAATGCAAGTGGCTTCAAAAAATGACAAATGAAATTCCTTTAAGAGACTTTGGT 1060
QY 960 TTAATGCAAGTGGCTTCAAAAAATGACAAATGAAATTCCTTTAAGAGACTTTGGT 1019
DB 1061 TTAATGCAAGTGGCTTCAAAAAATGACAAATGAAATTCCTTTAAGAGACTTTGGT 1120
QY 1020 ACTCCATTTGAAACATTAATCAGAGAGATTAAGCTAATTCCTTTAAGAGACTTTGGT 1079
DB 1121 ACTCCATTTGAAACATTAATCAGAGAGATTAAGCTAATTCCTTTAAGAGACTTTGGT 1180
QY 1080 AATATTTGAAAGCCCTGCTTAAAGAGATTCCTTCCGCGGCTTAAATTTGGG 1139
DB 1181 AATATTTGAAAGCCCTGCTTAAAGAGATTCCTTCCGCGGCTTAAATTTGGG 1240
QY 1140 ACTTGGGTGTAGCAACAGTGTGAGAGTCTTTCGTTGAGAGATTAAGCTGAGAGAG 1199
DB 1241 ACTTGGGTGTAGCAACAGTGTGAGAGTCTTTCGTTGAGAGATTAAGCTGAGAGAG 1300
QY 1200 AAGGCCCCGGGCTATCTTATGATGCTTATTTGTTGGGCTTGGGAGATTAAGAGCCAG 1259
DB 1301 AAGGCCCCGGGCTATCTTATGATGCTTATTTGTTGGGCTTGGGAGATTAAGAGCCAG 1360
QY 1260 CGGAGTAAATTAACCCCGCAGTGGCTTGGCTCCCTGAGAGCCCAATGAGTCAAGAA 1319
DB 1361 CGGAGTAAATTAACCCCGCAGTGGCTTGGCTCCCTGAGAGCCCAATGAGTCAAGAA 1420
QY 1320 TGTCCCAAGACTTGGCTTCCCAAGAGATTAATTTAGAAAGCTTCAAGAGTACAC 1379
DB 1421 TGTCCCAAGACTTGGCTTCCCAAGAGATTAATTTAGAAAGCTTCAAGAGTACAC 1480
QY 1380 GAATGGGAGCCCTGTATGAGAGATTCCTTCCGAGACAGGCTTCCGAGAGAGAGCTCAGC 1439
DB 1481 GAATGGGAGCCCTGTATGAGAGATTCCTTCCGAGACAGGCTTCCGAGAGAGAGCTCAGC 1540
QY 1440 CCGAGAGCGAAGGGGTAAAGGCGCGGGGTCAAGAGATTCGGGGGTGCTGAGTTGGCA 1499
DB 1541 CCGAGAGCGAAGGGGTAAAGGCGCGGGGTCAAGAGATTCGGGGGTGCTGAGTTGGCA 1600
QY 1500 GGAATGACTGGGGTGAACGGGGGTGCTGAGTGGCTGAGTGGCGGGGTGGCCGGGAGC 1559
DB 1601 GGAATGACTGGGGTGAACGGGGGTGCTGAGTGGCTGAGTGGCGGGGTGGCCGGGAGC 1660
QY 1560 ACCTTGGTTCTTGTAGAGAGAGTGAAGGGCTCCGGGCTGAGAGAGAGAGAGAG 1619
DB 1661 ACCTTGGTTCTTGTAGAGAGAGTGAAGGGCTCCGGGCTGAGAGAGAGAGAGAGAG 1720
QY 1620 CCGGGGCGAGTGTCAAGTGTCAAGGCGCGGCTGAGAGAGAGAGAGAGAGAGAGAG 1679
DB 1721 CCGGGGCGAGTGTCAAGTGTCAAGGCGCGGCTGAGAGAGAGAGAGAGAGAGAGAG 1780
QY 1680 CGGGCCCGGAGCGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739
DB 1781 CGGGCCCGGAGCGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840
QY 1740 GCGTTCCGGGGGTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
DB 1841 GCGTTCCGGGGGTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
QY 1800 CGGCCAG 1806
DB 1901 CGGCCAG 1907

RESULT 3
LOCUS HUMIBP3 10884 bp DNA linear PRI 04-APR-2002
DEFINITION Homo sapiens growth factor-binding protein-3 precursor (IGFBP3)
ACCESSION M35878 J05537 J05538 M35879 M35880 M35881 M35882 M35883 M35884
M35885 M35886 M36121 M36122
VERSION M35878.1 GI:184522
KEYWORDS Insulin-like growth factor binding protein 3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 10884)
AUTHORS Cubbage,M.L., Suwanichkul,A. and Powell,D.R.
TITLE Insulin-like growth factor binding protein-3. Organization of the
human chromosomal gene and demonstration of promoter activity
JOURNAL J. Biol. Chem. 265 (21), 12642-12649 (1990)
MEDLINE 90324259
PubMed 1695633
COMMENT On Apr 4, 2002 this sequence version replaced g1:763445.
Draft entry and computer-readable sequence kindly submitted by
D.R.Powell, 03-JUL-1990.
FEATURES
Source Location/Qualifiers
1..10884
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HL1006d"
/cell_type="Leukocyte"
102..10775
/gene="IGFBP3"
102..1905
/gene="IGFBP3"
/note="G00-126-724"
1808..1821
/gene="IGFBP3"
/note="G00-126-724"
1876..1881
/bound_molety="Spl and Ab-2"
/gene="IGFBP3"
/note="G00-126-724"
1876..1881
/gene="IGFBP3"
/note="G00-126-724"
1906..2440
/note="G00-126-724"
/product="growth factor-binding protein-3"
/gene="IGFBP3"
1906..2440
/gene="IGFBP3"
/note="G00-126-724"
/number=1
join(2038..2440,5726..5952,6497..6616,8212..8337)
/gene="IGFBP3"
/note="The AAs encoded by bases 5728-5736 and 5917-5926
may be Asn-linked glycosylation sites; Insulin-like
precursor"
/codon_start=1
/product="growth factor-binding protein-3 precursor"
/protein_id="AA52706.1"
/db_xref="GI:386791"
/db_xref="GDB:G00-126-724"
/translation="MORARPTMAAALTLVLRGPVVARAGASGGIGPVYRCPCD
ARALACAPPAVCAELVREPCGCCCLCALSEGPCGCTPRCCSGLRQCSPPDEAR
PLQALDGRGICLVNANSAVSRRLAYLLPAPAPAGNASESEEDSAGSVEPSSTHRY
SDPKFHLPLHSK111IKGHANDSQRYKVDYESQYDTQNFSSSESRRETEYGCRRME
DTLNLKFLNLVSPRGVH1PNDCKRGFKKQCRPSKGRKRGFCWCVDKYGQPLPGYT
TKGEDVHCYSNMQSK"
2038..2118
/gene="IGFBP3"
/note="G00-126-724"
join(2119..2440,5726..5952,6497..6616,8212..8334)
/gene="IGFBP3"
/product="growth factor-binding protein-3"
/note="G00-126-724"
5726..5952
/gene="IGFBP3"
exon

exon /note="G00-126-724"
/number=2
6497..6616
/gene="IGFBP3"
/note="G00-126-724"
/number=3
8212..10775
/gene="IGFBP3"
/note="G00-126-724"
/number=4
10751..10756
/gene="IGFBP3"
/note="G00-126-724"
polyA_signal
BASE COUNT 2796 a 2578 c 2737 g 2773 t
ORIGIN
Query Match 99.4%; Score 1795; DB 9; Length 10884;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGGATTCGTTTGTTCCTTCATTTTCCAAATGAAATCAGAGATCTGTTGGTGT 60
DB 101 GGGGATTCGTTTGTTCCTTCATTTTCCAAATGAAATCAGAGATCTGTTGGTGT 160
QY 61 CAACGAGATCTAGAGAGAGGTGATACAGAGAAAG-AACGACAGCGATTTATG 119
DB 161 CAACGAGATCTAGAGAGAGGTGATACAGAGAAAGAAACAGACAGATTTATG 220
QY 120 CACGTTCTCTGAACAAGGTGAGTACGACACCTGACACTGAGGAGAAAGC 179
DB 221 CACGTTCTCTGAACAAGGTGAGTACGACACCTGACACTGAGGAGAAAGC 280
QY 180 TCATAAGAAATAGCGTCTGGGCTTGCACCCGGGCGCTCAATGTTCTTGTCT 239
DB 281 TCATAAGAAATAGCGTCTGGGCTTGCACCCGGGCGCTCAATGTTCTTGTCT 340
QY 240 TTGGCTCTTTTATTTGTAGAGGTCCAAATTTATTTATTTAGTACAGGAGACA 299
DB 341 TTGGCTCTTTTATTTGTAGAGGTCCAAATTTATTTATTTAGTACAGGAGACA 400
QY 300 AATTGATCTTCATCTAAAGAGAGATATATGATTAAGAGAGCTGATATAT 359
DB 401 AATTGATCTTCATCTAAAGAGAGATATATGATTAAGAGAGCTGATATAT 460
QY 360 GGGGAGAGAGGTGACAGGGGGGAAAGGGAGAGACGAGAGAGAAAGGAGGAGAG 419
DB 461 GGGGAGAGAGGTGACAGGGGGGAAAGGGAGAGACGAGAGAGAAAGGAGGAGAG 520
QY 420 GGCAGAGAGAGACTGGCGGAGAGATGATTAGAGAGACAGAAATGATGATGAGA 479
DB 521 GGCAGAGAGAGACTGGCGGAGAGATGATTAGAGAGACAGAAATGATGATGAGA 580
QY 480 TTAACTTCAACCAGGCTTCGTCGTGAGGGGAATGAGGAGCTCGATTGCTATTA 539
DB 581 TTAACTTCAACCAGGCTTCGTCGTGAGGGGAATGAGGAGCTCGATTGCTATTA 640
QY 540 CTACTCCAAACTGCAAGGGGCTCTCAAGTCACTATCCACTCTTAAGCAAGCTGCC 599
DB 641 CTACTCCAAACTGCAAGGGGCTCTCAAGTCACTATCCACTCTTAAGCAAGCTGCC 700
QY 600 AATTTCACAGCGTTCAGAGAAAGTCTCCCGGAGAGTCTACCGCTTCCACTCAG 659
DB 701 AATTTCACAGCGTTCAGAGAAAGTCTCCCGGAGAGTCTACCGCTTCCACTCAG 760
QY 660 CCCCAAAACTCTTGGAAAAGTCTTAAATAATTTATCTCAATCTCTGAGAC 719
DB 761 CCCCAAAACTCTTGGAAAAGTCTTAAATAATTTATCTCAATCTCTGAGAC 820
QY 720 ACCAGGCTCTCTGTTGTCACCGAAGAGGGGTGCGAGACAAATGGAAGAACTCG 779
DB 821 ACCAGGCTCTCTGTTGTCACCGAAGAGGGGTGCGAGACAAATGGAAGAACTCG 880
QY 780 AGTGCAGAGAAAGCCGACAGAGATTACAGCACTCAGCGGCAATTGCCGCCGAAC 839

```

|||||
Db 881 AGTGCAGGAAAGCGACAGGAGTTACAGCCACTCAGCGGCAATTCGCCCGAACT 940
Oy 840 TTACTGAAAGTGTAGATTGACAGATAGCTAGAAATCCACAGCATCAGAAATACAG 899
Db 941 TTACTGAAAGTGTAGATTGACAGATAGCTAGAAATCCACAGCATCAGAAATACAG 1000
Oy 900 TAATACGAAGTGCCTTCAAAAATGACAAATGCAATGCAATGCTATTAAAGACTATTGGT 959
Db 1001 TAATACGAAGTGCCTTCAAAAATGACAAATGCAATGCAATGCTATTAAAGACTATTGGT 1060
Oy 960 TAATACGTTTACAGAGTGCCTTCAAAAATGCAATGCAATGCTATTAAAGACTATTGGT 1019
Db 1061 TAATACGTTTACAGAGTGCCTTCAAAAATGCAATGCAATGCTATTAAAGACTATTGGT 1120
Oy 1020 ACTCATTTGAAATATATATATGAGGAGATATCCCAAGACAAAGAGAGCTGTCAATTA 1079
Db 1121 ACTCATTTGAAATATATATGAGGAGATATCCCAAGACAAAGAGAGCTGTCAATTA 1180
Oy 1080 AAATATTTGAAAGCCCTGCTTAAAGAGCATTCCTGCTGCTCACTTAAATTTGGG 1139
Db 1181 AAATATTTGAAAGCCCTGCTTAAAGAGCATTCCTGCTGCTCACTTAAATTTGGG 1240
Oy 1140 ACTTCGGGTGTAGCAACAGTGAAGTCTTCTGCGTTGAGAGTAAAGCTTGAAGAGCG 1199
Db 1241 ACTTCGGGTGTAGCAACAGTGAAGTCTTCTGCGTTGAGAGTAAAGCTTGAAGAGCG 1300
Oy 1200 AAGGCCCGGGGCACTTCAGATGCTATTTGTGGGCCCCCTGGGATATAAACAGCCAG 1259
Db 1301 AAGGCCCGGGGCACTTCAGATGCTATTTGTGGGCCCCCTGGGATATAAACAGCCAG 1360
Oy 1260 CGGGTGAATTAACCCCGAGTGCCTTGGCTGCTGAGACCCAAATGTAGTACAGAA 1319
Db 1361 CGGGTGAATTAACCCCGAGTGCCTTGGCTGCTGAGACCCAAATGTAGTACAGAA 1420
Oy 1320 TGTCCCAAGACTTGGCTGCTGAGAGTAAATTTAGAAAGCTCCAGAGAGTACACAC 1379
Db 1421 TGTCCCAAGACTTGGCTGCTGAGAGTAAATTTAGAAAGCTCCAGAGAGTACACAC 1480
Oy 1380 GAATCGGAGCGCTGTATAGCAGATTTCCCGACACCGGCTGCGCAGGAGAGACTCAC 1439
Db 1481 GAATCGGAGCGCTGTATAGCAGATTTCCCGACACCGGCTGCGCAGGAGAGACTCAC 1540
Oy 1440 CCGAGAGCGGAAGGGGTAAAGGGGCGGAGTCAAGAGATCGGGGGTGTGAGTTGGCCA 1499
Db 1541 CCGAGAGCGGAAGGGGTAAAGGGGCGGAGTCAAGAGATCGGGGGTGTGAGTTGGCCA 1600
Oy 1500 GGAAGTACTGGGGGTACCGGGGGTGTCTAGAGTGGCTGGAGTCCCGGGGTGCGCGGAC 1559
Db 1601 GGAAGTACTGGGGGTACCGGGGGTGTCTAGAGTGGCTGGAGTCCCGGGGTGCGCGGAC 1660
Oy 1560 ACCTTGGTCTTGTAGACAGACAAAGGTGAGCGGGCTCCGGGCGTGAACAGAGCAGGTG 1619
Db 1661 ACCTTGGTCTTGTAGACAGACAAAGGTGAGCGGGCTCCGGGCGTGAACAGAGCAGGTG 1720
Oy 1620 CCGGGGCGAGTCTGAGTGCACGCGCCCGAGCTGCGCGCGGCTCTAGGGCGAAGCA 1679
Db 1721 CCGGGGCGAGTCTGAGTGCACGCGCCCGAGCTGCGCGCGGCTCTAGGGCGAAGCA 1780
Oy 1680 CCGGGCGGCGAGCGTGGCTGCGCGGACCGCGCCCTCCCAACCCCACTCTCTGGGCGC 1739
Db 1781 CCGGGCGGCGAGCGTGGCTGCGCGGACCGCGCCCTCCCAACCCCACTCTCTGGGCGC 1840
Oy 1740 GCGTTCGCGGGGCTGCTGCGGGGACCGCGGCTCTATATACGGGCGGCGCGGCGG 1799
Db 1841 GCGTTCGCGGGGCTGCTGCGGGGACCGCGGCTCTATATACGGGCGGCGCGGCGG 1900
Oy 1800 GCGCCAG 1806
Db 1901 GCGCCAG 1907

```

RESULT 4

```

AC091524/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

AC091524 69887 bp DNA linear PRI 07-NOV-2001
Homo sapiens BAC clone RP11-132L11 from 7, complete sequence.
AC091524
AC091524.4 GI:16756331
HTG.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 69887)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
9847074
2 (bases 1 to 69887)
Mulvaney, E. and Haakenson, W.
The sequence of Homo sapiens BAC clone RP11-132L11
Unpublished
3 (bases 1 to 69887)
Waterston, R.H.
Direct Submission
Submitted (29-APR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 69887)
Waterston, R.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 7, 2001 this sequence version replaced gi:14029081.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.edu
----- Summary Statistics
Center project name: H_NH0132L11

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send
mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPc11 human BAC library was made from the blood of one male
donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J., and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is Rp11-63818; the clone sequenced to the right is Rp11-571N3, 200 bp overlap. Actual start of this clone is at base position 1 of Rp11-132L11; actual end is at base position 120387 of Rp11-571N3.

There are polymorphic base pair differences in the overlap between the clone Rp11-132L11, Rp11-63818, Rp11-571N3 and redundant clone Rp11-14802.

```
FEATURES
Source
Location/Qualifiers
1..69887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="Rp11-132L11"
/clone_id="RPI1-11"
367..464
/rpl_family="MIRL_type"
/rpl_family="MIR"
1326..1440
/rpl_family="MIR"
2228..2260
/rpl_family="(CA)n"
2462..2467
/note="match to EST AA368830 (NID:g2021148)"
2913..3072
/rpl_family="CT-rich"
3047..3354
/rpl_family="Alu"
4681..4752
/rpl_family="L2"
5158..5180
/rpl_family="AT-rich"
5161..5462
/rpl_family="Alu"
5670..5887
/rpl_family="L1"
6217..6356
/rpl_family="MIR"
7456..7958
/note="match to EST BG695859 (NID:g13960406)"
7461..7958
/note="match to EST AL551522 (NID:g12889551)"
7462..7958
/note="match to EST AL551522 (NID:g12889551)"
7462..7958
/note="similar to Homo sapiens EST AL553038 (NID:g12892497)"
7462..7825
/note="match to EST AL551976 (NID:g12890432)"
7463..7958
/note="match to EST AL553501 (NID:g12893394)"
7467..7919
/note="match to EST AL570357 (NID:g12926584)"
7470..7958
/note="similar to Homo sapiens EST AL552017 (NID:g12890512)"
7486..7958
/note="match to EST AL531252 (NID:g12794745)"
7548..7924
/note="match to Bos taurus EST BG666453 (NID:g10027044)"
7802..7958
/note="similar to Mus musculus EST A1892189 (NID:g5598091) mJ88b01.y1"
7807..7958
/note="similar to Homo sapiens EST AL545650 (NID:g12878172)"
7812..7958
/note="match to EST AV650862 (NID:g9871876)"
7816..7958
/note="match to EST AL570706 (NID:g12927278)"
7845..7958
/note="similar to Homo sapiens EST AL571370 (NID:g12928599)"
```

```
misc_feature 7849..7958
/note="match to EST AL577394 (NID:g12940479)"
misc_feature 7862..7958
/note="match to EST AA368988 (NID:g2021308)"
misc_feature 7880..7958
/note="similar to Homo sapiens EST AL571308 (NID:g12928474)"
misc_feature 7931..7958
/note="match to EST AL571695 (NID:g12929247)"
repeat_region 7969..8031
/rpl_family="C-rich"
7985..7990
/note="match to EST AL569218 (NID:g12924337)"
8513..8676
/note="match to EST AA077843 (NID:g1837317)"
8551..8570
/note="match to EST BG422582 (NID:g13329088)"
8596..9243
/note="match to EST AV685060 (NID:g10286923)"
8596..9195
/note="match to EST AV692462 (NID:g10294325)"
8807..9098
/note="match to EST AV687868 (NID:g10289731)"
9103..9369
/note="match to EST AV659647 (NID:g9880661)"
9154..9241
/note="match to EST AV687868 (NID:g10289731)"
9155..9156
/note="match to EST W01199 (NID:g1272178) za59a07.r1"
9172..9457
/note="match to EST T46868 (NID:g648854) yb94d09.r1"
9296..9718
/note="match to EST AW962557 (NID:g8152393)"
9424..9519
/note="match to EST AV659647 (NID:g9880661)"
9473..9674
/note="similar to Mus musculus EST A1892189 (NID:g5598091) mJ88b01.y1"
9505..9906
/note="match to EST AL531252 (NID:g12794745)"
9505..9674
/note="match to EST AL551522 (NID:g12889551)"
9505..9674
/note="match to EST AL553501 (NID:g12893394)"
9505..9674
/note="match to EST AL570706 (NID:g12927278)"
9505..9674
/note="match to EST AL571695 (NID:g12929247)"
9505..9674
/note="match to EST AL577394 (NID:g12940479)"
9505..9674
/note="match to EST AV650862 (NID:g9871876)"
9505..9674
/note="match to EST BG695859 (NID:g13960406)"
9505..9674
Query Match 93.3% Score 1685.4; DB 9; Length 69887;
Best Local Similarity 98.7% Pred. No. 0;
Matches 1794; Conservative 0; Mismatches 11; Indels 12; Gaps 9;
```

||||| 41863 TCATAGAAAAAGAGGCTGCTGGCCCTGTCACCCCGGGGACATCCATGTTCTTGTC 41804
QY 240 TTGGCTCTCTTTTATTTAGAGTCCATATTTATTTATTTAGTACAGAGGAGCA 299
41803 TTGGCTCTCTTTTATTTAGAGTCCATATTTATTTATTTAGTACAGAGGAGCA 41744
QY 300 AATTGATCTTTCCATTTTAAAGAGAGTATATATGTATTAAGAGAGCTGTATAGAT 359
41743 AATTGATCTTTCCATTTTAAAGAGAGTATATATGTATTAAGAGAGCTGTATAGAT 41684
QY 360 GGGGGAAGAGTGGACAGGGGGGAAAAAGGAGAGACGACAGAGAGAAAGGAGAGAG 419
41683 GGGGGAAGAGTGGACAGGGGGGAAAAAGGAGAGACGACAGAGAGAAAGGAGAGAG 41624
QY 420 GCACAAAGAGAGACACTGGGCGAGAGATGATTAGAGAGACAGAAATATGATGATGATA 479
41623 GCACAAAGAGAGACACTGGGCGAGAGATGATTAGAGAGACAGAAATATGATGATGATA 41564
QY 480 TTAACTTCAACCAAGGCTCTGCTGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 539
41563 TTAACTTCAACCAAGGCTCTGCTGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 41505
QY 540 CTAACTTCAACCAAGGCTCTGCTGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 599
41504 CTAACTTCAACCAAGGCTCTGCTGAGAGGGAATGAGAGAGCTCTGATTTGCTATTA 41445
QY 600 AATTTCACAGGCTTCAGAGAAAGTCTCTCCGCGAGAGTCTACCGTCTCCACATCCAC 659
41444 AATTTCACAGGCTTCAGAGAAAGTCTCTCCGCGAGAGTCTACCGTCTCCACATCCAC 41385
QY 660 CCCCAACAACTCTTGGAAAAAGTGCCTTGAATAATTTAACTCAATCCATCTGAGAC 719
41384 CCCCAACAACTCTTGGAAAAAGTGCCTTGAATAATTTAACTCAATCCATCTGAGAC 41325
QY 720 ACCAGGCTCTCTGTTGGTGCACCGAA - GGAAGGGGTGGCCACACAAATGAAAGAAC 778
41324 ACCAGGCTCTCTGTTGGTGCACCGAAAGGAGGGGTGGCCACACAAATGAAAGAAC 41265
QY 779 GAGTCCAGAGAGAGCCGACAGAGATTACAGGACCTCAGCCGCAATTCG - CCCCAGAA 837
41264 GAGTCCAGAGAGAGCCGACAGAGATTACAGGACCTCAGCCGCAATTCG - CCCCAGAA 41205
QY 838 CTTTACTGAAAAAGTGTATGATTCAGAGATTAAGCTAAGATCCCAAGCATGAGAAATAC 897
41204 CTTTACTGAAAAAGTGTATGATTCAGAGATTAAGCTAAGATCCCAAGCATGAGAAATAC 41145
QY 898 ACTAATAGCAAGTGGCTCAAAAAATGCAANTGAATAATGGCTAATTAAGCACTATTG 957
41144 ACTAATAGCAAGTGGCTCAAAAAATGCAANTGAATAATGGCTAATTAAGCACTATTG 41085
QY 958 GTTAAATTACGTTTACAGCACTGCCAGATTATTTGTTTATTTATTTTGTGCGTGGTGT 1017
41084 GTTAAATTACGTTTACAGCACTGCCAGATTATTTGTTTATTTATTTTGTGCGTGGTGT 41025
QY 1018 AAACCTCAATTTGAAAAATTAATCAGGAGAAATACCAAGCAAGAAACAGTTGTCAAT 1077
41024 AAACCTCAATTTGAAAAATTAATCAGGAGAAATACCAAGCAAGAAACAGTTGTCAAT 40965
QY 1078 TAAATATTTGAAAAAGCCCTGCTTAAGGA - GCATTGCTGTCGCGGTCCACCTTAAATG 1136
40964 TAAATATTTGAAAAAGCCCTGCTTAAGGAGCATTCGCTGCGGTCCACCTTAAATG 40905
QY 1137 GGGACTTGGGTGTAGCAACAGTGAAGTCTTCTGCTTGAAGTGAAGTGAAGTGAAG 1196
40904 GAGACTTGGGTGTAGCAACAGTGAAGTCTTCTGCTTGAAGTGAAGTGAAGTGAAG 40845
QY 1197 GCG - - - AAGCCCGCGGCGCATCTTCAAGTGCATTTGTTGGCCCGCTGGGATATAACA 1253
40844 GGGCGAGAGCGCGCGCGCATCTTCAAGTGCATTTGTTGGCCCGCTGGGATATAACA 40785
QY 1254 GCGCAGCGGCTGTAATTAACCCGCGCATGCTTGGTCCCGGAGACCCCAATGTAAAG 1313
|||||

Db 40784 GCGCAGCGGCTGTAATTAACCCGCGCATGCTTGGCTCCCTGAGACCCCAATGTAACT 40725
QY 1314 CAGAAATGTCCCAAGACTTCCGCTGCCAACGAAATTAATTTAGAAAGCTCCACGAGT 1373
Db 40724 CAGAAATGTCCCAAGACTTCCGCTGCCAACGAAATTAATTTAGAAAGCTCCACGAGT 40665
QY 1374 ACACAGCAATGCGGAGCGCTGTATGCCAGTTTCCCGACACCGGCTCGCCGAGGAGAC 1433
40664 ACACAGCAATGCGGAGCGCTGTATGCCAGTTTCCCGACACCGGCTCGCCGAGGAGAC 40605
QY 1434 CTCACCCCGAGACCGGAAAGGGGTAAAGGGCGGGGGGCAAGAGATCGGGGGTGTAGT 1493
40604 CTCACCCCGAGACCGGAAAGGGGTAAAGGGCGGGGGTCAAGAGATCGGGGGTGTAGT 40545
QY 1494 TGGCCAGAGTGAAGTGGGTGACCGGGGGGTCTGAGGTGCTGAGTGGCTGGGGGTGGCC 1553
40544 TGGCCAGAGTGAAGTGGGTGACCGGGGGGTCTGAGGTGCTGAGTGGCTGGGGGTGGCC 40485
QY 1554 GGGCACAACCTTGGTCTTGTAGACGACAAAGTGA - - CGGGCTCCGGGCGTGAACGAGG 1611
40484 GGGCACAACCTTGGTCTTGTAGACGACAAAGTGAAGTGAACCGGGGCTCGGGCGTGAACGAGG 40425
QY 1612 AGCAGGTGCGCGGGCGAGTCTGAGGTGACAGGCGCCCGAGCTGGGCGCGGCTGAGG 1671
40424 AGCAGGTGCGCGGGCGAGTCTGAGGTGACAGGCGCCCGAGCTGGGCGGCTGAGG 40365
QY 1672 GCGAAGCAGAGG - CCGCGAGCGGCTGCGGCGAGACCGCCCGCTCCCAACCCCACT 1730
40364 GCGAAGCAGAGGCGCGCGAGCGGCTGCGGCGAGACCGCCCGCTCCCAACCCCACT 40305
QY 1731 CTTGGGCGCG - CGTTCGGGGGCGTGTCTGGGCGACCCCGGCTTATATACGGGCGGC 1789
40304 CTTGGGCGCGCGCTTCCGGGGCGTGTCTGGGCGACCCCGGCTTATATAGGCGCGGC 40245
QY 1790 GCGCGCGGGCGCGCGAG 1806
40244 GCGCGCGGGCGCGCGAG 40228
RESULT 5
AX345904 8700 bp DNA linear PAT 01-FEB-2002
LOCUS AX345904
DEFINITION Sequence 975 from Patent WO0200928.
ACCESSION AX345904
VERSION AX345904.1 GI:18493790
KEYWORDS
SOURCE
ORGANISM synthetic construct.
synthetic construct.
artificial sequences.
REFERENCE
1 Olek, A., Plepenbrock, C. and Berlin, K.
Diagnosis of diseases associated with the immune system
Patent: WO 0200928-A 975 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
source location/Qualifiers
1..8700
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 2247 a 281 c 2233 g 3939 t
ORIGIN
Query Match 63.2%; Score 1141.4; DB 6; Length 8700;
Best Local Similarity 80.0%; Pred. No. 1.7e-260;
Matches 1454; Conservative 0; Mismatches 351; Indels 12; Gaps 9;
QY 1 GGGGATTCGTTTGTCTTCCATTTTCCAAATGAATCAGAGATCCTGTTTGGGTGT 60
Db 3055 GGGGATTCGTTTGTCTTGTATTTTAAATTTTAAATTTGAGATTTGTTTGGGTGT 3114
QY 61 CACGCGAGATCTAGAGAGAGTGTATACAGAGAAAGG - AACAGCAAGCAGCATTTATG 119
Db 3115 TAACTGATATTAGAGAGAGTGTATTAAGAGAAAGAAATGTATAGCAGCATTTATG 3174
|||||

OY	120	CACGCTTCCTGTAACAGAGTTGAGTGAAGCCACACCTGACACTGAGGGGGAAGAC	179
Db	3175	TACGGTTTTTGTAAATRAAGTTGAGTGTACTTATGATTGATATTGGGAGAAAGAT	3233
OY	180	TCATAGAAAATGACGGTGTGGGCTTCGTCAACCCGGGGCCCTCCATTGTTCTGTCT	239
Db	3235	TTATAGAAAATGACGGTGTGGGCTTCGTCAATTCGGGGATTTTAAATGTTTTGTTT	3294
OY	240	TTGGCTCTTTTATTTGTAGAGTCCAAATATTTATTTTATAGTACAAAGGGAGCA	299
Db	3235	TTGGTTTTTTTTTATTTGTAGAGTTCATATTTATTTATTTATTTATTTAGTAAAGGGAGCA	3355
OY	300	AATGATCTTTCCATTCCAAAAGAGAGTATATGTAATTAAGAGAAAGCTGTATGATAT	359
Db	3355	AATGATTTTTTTTATTTTAAAAGAGATATATGTAATTAAGAGAAAGTGTATGATAT	3411
OY	360	GGGGAGAGAGTGCACAGAGGGGAAAAGGGAGAGAGACAGAGAGAAAAGGAGAGAG	419
Db	3415	GGGGAGAGAGTGTATAGGGGGAAAAGGGAGAGAGACAGAGAGAAAAGGGAGAGAG	3477
OY	420	GGACAGAGAGAGACAGTGGGCGAGATCGATTAGAGAGACAGAAATGATGATGAAGA	479
Db	3475	GGATRAGAGAGATRTTGGCGAGATCCATTAGAGAGATGAAATGATGATGAAGA	3533
OY	480	TTAATCTCACCAAGCTTCCTCGCTGGAGGGAGATGAGAGACTCGATTTGCTATTA	539
Db	3535	TTAATTTTTTTTAAAGTTTCTGTGTGAGAGGGAAATGAGAGATTTTGATTTATTA	3595
OY	540	CTACTCCAAATCGAAAAGGCTCCTTCAATCACTATCCACTCCCTAAGGCAAGCTCC	599
Db	3534	TTATTTTAAATTTGTAAGGGTTTTTTTAACTATTTATTTATTTTAAAGTAAAGGCTTT	3655
OY	600	AATTTCAACAGCGTTTCAGGAAAGTCTCTCCCGGAGGTCTACCGCTCTCCACTCCAC	659
Db	3654	AATTTTAAATGCTTTTAAAGAAAGTTTTTTTCCGGAGGTTTTATGTTTTTATTTAT	3711
OY	660	CCCCAACAACCTTTGGAAAAAGTCCCTGAAAAATTTAATCCATATCCATCTGGACC	719
Db	3714	TTTTTAAATTTTTTGGAAAAAGTCTTTGAAAAATTTAATTTAATTTAATTTGGATT	3773
OY	720	ACCAAGCTCCTCTGTGGTCAACGAA-GGAGGGGGCCGACAGCAAAACGTGAAGAACTC	778
Db	3774	ATTAGCGTTTTTGTGGTATTCGAAGGAGGGGGTCCGTAGATTAATTTGAAGAAATTC	3833
OY	779	GAGTGCAGAGAGAGCCGACAGAGATTACAGCACTCAAGCGCGCAATTGCG-CCCCGAA	837
Db	3834	GAGCTGTAGAGAAAGTGTAGAGATTAGCCATTTTGAACCGGTATTTGCGTTTGCAGAA	3893
OY	838	CTTTACGAAAAGTGTAGATTGCACAGATTAAGCTAGAAATCCCAAGCATGAGAGATAC	897
Db	3894	TTTTTATGAAAAGTGTAGATTGTGTAGATTAAGATTTTAAAGTATGAGAAATAT	3955
OY	898	AGTAAATAGAGATCGCTTCAAAAAGACAAATGAATAATGCTATTTAAGACTATTTG	957
Db	3954	AGTAAATAGAGATCGCTTTTAAAAAATGATAAGAAATGCTTATTTAAGATTTATGG	4013
OY	958	GTTAATTAACGTTTACAGAGTCCCAAGTTATGTCTTATTTATTTCTTTGTGCTGGGT	1017
Db	4014	GTTAATTAACGTTTATGATAGTGTAGTTATTTGTATTTTATTTATTTTGTGCTGGGT	4073
OY	1018	AAATCTCAATTTGAAGAAATATATCAGGAGAAATACCAAGACAAGAGAACAGTGTGCAT	1077
Db	4074	AAATTTTATTTGAAGAAATATATTTAGGAGAAATTTTAAAGTAAAGAAATATGTTGTAT	4133
OY	1078	TAAAAATATTTGAAAAGCCCTGCTTTAAGGA-GCATTTGCTTGGCGGTCCACTTTAATG	1136
Db	4134	TAAAAATATTTGAAAAGTGTGTTTAAAGCACTATTCGTTTGGCGGTATTTAATTTG	4193
OY	1137	GGGACTTGGCGGTAGAGACAACGAGAGACTCTTGCGGTAGAGTAAAGCTGGAAAG	1199
Db	4194	GAGATTTGGCGGTAGATTAATRCGAGATTTTTTTTTTGGCGGTGAAGTAAAGTTTGAAG	4255

OY	1197	GC-----AAGGCCCGGGGCATCTTCGAGATGCGCATTTTGGGCCCCGGGATTTAAACA	1253
Db	4254	GGCGGAAAGGTTTCGGCGGATTTTTCGATGACGATTTTGGGTTTTCGGGATTTAAATA	4313
OY	1254	GCCGACGGGGGTGTAATTTAAACCCGCGAGTGCCTTGGCTCCCTGAGACCCAAATGTAAGT	1313
Db	4314	GTTTACGGGGGTAAATTAATTTCCGTAAGTGGTTTGGTTTTCGATTTTAAATGTAAGT	4373
OY	1314	CAGAAATGTCOCGAAGTTCGCTCGCCAAAGGAAATTAATTTTGAAGAAAGTCCACGAGGT	1373
Db	4374	TAGAAATGTTTAAAGTTTCGTTTGTTCGGAATTTAAATTTTGAAGAGTTTACGAGGT	4433
OY	1374	ACACACAAATGCGGAGCGCGTGTATGCGAGTTTCCCGACACCGCGCTGCGCCGACGAGAC	1433
Db	4434	ATAATGAAATGCGGAGGGTGTATGTATGTTTTCGATATCGGTTGCTGTAGGAGAT	4493
OY	1434	CTCACCCCGAGAGCGGAAGGGGTAAAGGCGCGGGGTCAAGAGATCGGGGCTGTAGT	1493
Db	4494	TTTATTTTCGAGACGCGGAAGGGGTAAAGCGCGGGTTAAAGAAATCGGGGCTGTAGT	4553
OY	1494	TGGCGAAGATGACTGGGGTGACCGGGGGGTGTAGAGTGGCTGGACTGGCGGGGTGGCC	1553
Db	4554	TGGTAAAGATGATGGGGTGATCGGGGGTGTGTAGAGTGGTGAAGTGTGCGGGGTGGTC	4613
OY	1554	GGCGACACTCTGGTGTCTGTAGACGACAAAGTGTGA--CGGGCTCGGGCGTAGACAGAG	1611
Db	4614	GGGTATATTTTGGTTTTTGTAGACGATTAAGTGAATTCGGGTTTGGGGGTGCGTAGAG	4673
OY	1612	AGCAGGTGCCCCGGCGAGTCTGAGCTGCAGCGCCCCGACGACTCGGCCCCGGCTGCTAG	1671
Db	4674	AGTAAAGTGTTCGGGCGAGTTCGAGTTTCGACTGTTTTCGAGTTTCGAGTTTCGTTGTTAG	4733
OY	1672	GCGAAGACAGGAG--CCCGGACGCGGTCTCGGCGCGAACCCGCCCCCTCCCAACCCCACT	1730
Db	4734	GCGAAGACAGGTTTTTCGTAGTGTGTCTTTCGCTCAATTCGTTTTTTTAAATTTTATTT	4793
OY	1731	CTTGAGGCGCG--CGTTCGGGGCGGTGCTTGAGGCAACCCCGGCTTCTATATAGGCGCGC	1789
Db	4794	TTTGGGCGCGTCTGTTTCGCGGGCGTGTTTTGGGTATTTTCGTTTTTAAATAGCGTCGC	4853
OY	1790	GGCGCGGGGCGCGCCAG	1806
Db	4854	GCCTTCGGGCTGTTTAA	4870

AX345905/c	8700 bp	DNA	linear	PAT 01-FEB-2002
LOCUS	AX345905			
DEFINITION	Sequence 976 from Patent WO0200928.			
ACCESSION	AX345905			
VERSION	AX345905.1	GI:18493791		
KEYWORDS				
SOURCE	synthetic construct.			
ORGANISM	synthetic construct			
	artificial sequences.			
REFERENCE	1			
AUTHORS	Olek, A., Piepenbrock, C. and Berlin, K.			
TITLE	Diagnosis of diseases associated with the immune system			
JOURNAL	Patent: WO 0200928-A 976 03-JAN-2002;			
	Epigenomics AG (DE)			
FEATURES	Location/Qualifiers			
Source	1..8700			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="chemically treated genomic DNA (Homo sapiens)"			
BASE COUNT	2133 a	281 c	2087 g	4199 t
ORIGIN				
Query Match	56.5%;	Score 1021.2;	DB 6;	Length 8700;
Best Local Similarity	76.0%;	Pred. No. 6.6e-232;		
Matches 1377;	Conservative 0;	Mismatches 423;	Indels 12;	Gaps 9;
Y	5	ATTGCGTTTGTTTCCTTCACATTTCCAAATTCGAAATCCAGATCCGTGCTTCGGGTGCAAC	64	


```

||||| 5462 ATTCGTTATTTCTTCAATTTTCCAAATTAATCAAAAATCTATTTAAATATCAAC 5583
||||| 65 GCAGATCTAGAGAAGGATGATACAGAGAAGAA-CAGCAAGCCAGATTATGGCAG 123
||||| 5582 GCAAAATCTAAAAAATAATATCAAAAAAACAACAAGCAGATTATAGCAG 5523
||||| 124 GTTCTCTGTAACAAAGTTGAGTGACCAAGCTGAGCAGTGTGGAGAAAGCTCAT 183
||||| 5522 ATTTCTTAATAAACAATAATATTAACCACTAAACCTATTAATAAATAAATCAT 5463
||||| 184 AAGAAATAGAGGTGTGGCTGTGTCACCCGGGGCCCTCATGTCTGTCTTGG 243
||||| 5462 AAAAAATAGCATCTAAACCTGTGACCCGAAACCTCATTTATCTTTAA 5403
||||| 244 TCTCTTTTATTTAGAGGTCCAATTTATTTATTTAGTACAGAGGAGCAAAAT 303
||||| 5402 TCTCTTTTATTTAATAAATTCATTTATTTATTTAATCAAAAAAAGCAAAAT 5343
||||| 304 GATCTTTCTATTAAGAGAGATATATATTAAGAAAGCTGTATAGATATGGG 363
||||| 5342 AATCTTTCTATTAATAAATAATATATATATTAATAAATAAATATATATAA 5283
||||| 364 GAAGAGTGAGGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
||||| 5282 AAAAAATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5223
||||| 424 AAGAGAGAGAGAGAGAGAGAGATATAGAGAGAGAGAGAGAGAGAGAGAGAG 483
||||| 5222 AAAAAAACAATAAAGCAAAATATGATTAATAAATAAATAAATAAATAAATAA 5163
||||| 484 CTTCACCCAGAGCTTCGTGCTGAGAGGAGATGAGAGAGCTCTGATTTGCTATTA 543
||||| 5162 CTTCACCCAGAGCTTCGT-ACATAAATAAATAAATAAATAAATAAATAAATAA 5104
||||| 544 TCCAAAGTCAAGAGGCTCTCAATGTCATCTATCACTCTCAAGGAGAGGCTCAAT 603
||||| 5103 TCCAAAGTCAAGAGGCTCTCAATGTCATCTATCACTCTCAAGGAGAGGCTCAAT 5044
||||| 604 TCAACAGGCTTCAGAAAGTCTCTCCGCGAGAGTCTCACGCTCCCACTCAACCCC 663
||||| 5043 TCAACAGGCTTCAGAAAGTCTCTCCGCGAGAGTCTCACGCTCCCACTCAACCCC 4984
||||| 664 ACAAACTCTTGGAAAGTCTTGAATAATTAATCTCAATCTCAATCTGAGCAGCA 723
||||| 4983 ACAAACTCTTGGAAAGTCTTGAATAATTAATCTCAATCTCAATCTGAGCAGCA 4924
||||| 724 GGTCTCTCTGTGGTCAACG-AAGGAGGGGTGCGCAGCAAAATCTGAAGAAATCTGAGT 782
||||| 4923 AGTCTCTATTAATCAACGAAATAAATAAATAAATAAATAAATAAATAAATAA 4864
||||| 783 GCCAGAGAAGCGCAGAGAGATTTACAGAGCCTCAGCGCAATTGCG-CCCCGAATCTT 841
||||| 4863 ACCAAAAAAGCGCAAAATTTACAGAGCCTCAGCGCAATTGCGCCCCGAATCTT 4804
||||| 842 ACTGAAAGTGTGTTGATTTGAGAGATTAAGTAAATCCCAACGATCGAGATACAGTA 901
||||| 4803 ACTGAAAGTGTGTTGATTTGAGAGATTAAGTAAATCCCAACGATCGAGATACAGTA 4744
||||| 902 ATGAGAGGCGCTCAAAAAATGACAAATGCAATGCAATGCAATGCAATGCAAT 961
||||| 4743 ATGAGAGGCGCTCAAAAAATGACAAATGCAATGCAATGCAATGCAATGCAAT 4684
||||| 962 ATTAGCTTTACAGAGTGGCAGTATTTATTTATTTATTTATTTATTTATTTATTT 1021
||||| 4683 ATTAGCTTTACAGAGTGGCAGTATTTATTTATTTATTTATTTATTTATTTATTT 4624
||||| 1022 TCCATTTGAAAACATTAATCAGGAGATTAACCAAGCAAGAAAGAGTGTGCTATTAA 1081
||||| 4623 TCCATTTGAAAACATTAATCAGGAGATTAACCAAGCAAGAAAGAGTGTGCTATTAA 4564
||||| 1082 AATTTGAAAAGCGCTGCTTAAGA-GCATTCGCTTGGCGCTTCACTTTATTTGGGA 1140
|||||

```

```

Db 4563 AATTTAAAAAACCTTACTTAAAAAGCATTTGGCTTACGATCCACTTTAATTAATA 4504
Oy 1141 CTTCGGGTGAGACACAGTGAGAGTCTTCTGCTTGAAGAAAGCCT---GGAAGG 1197
||||| 4503 CTTCAGATTAATCAACAGCAAGTAAATCTTCTTACGTTTAAAAAATTAACCTAAAAA 4444
Oy 1198 CGAAGGCCCGGGGCACTTTCAGATCGTATTTTGTGGCCCTTGGGAGATTAACAGCC 1257
||||| 4443 GAAAGCCCGAAGCAGCTCTTCAATAGTATTTATTAACCCCTAAAAAATTAACAAACC 4384
Oy 1258 AGCGGTGAATTAACCCCGAGTCCCTGGCTCCCTGAGACCCAAATGAAGTCAG 1317
||||| 4383 AACGAATATTAATTAACCCCGAATATCTTAATCTCTTAACCCCAATATTAATTAAC 4324
Oy 1318 AATGTCCCAAGACTTTCGCTGCGCAAGCAATTAATTTAAGAAAGTCCAGAGTATAC 1377
||||| 4323 AATATCCCAAACTTTCGCTTACCAAGAAATTAATTTAAAAAATCCAGAAATATAC 4264
Oy 1378 AGAATGCGAGCGCTGTATGCCAGTTCCTCCGACACCGGCTCGCGAGAGAGACTCA 1437
||||| 4263 AGGAATACGAAGCGTATATACCAATTTCCCGACACCGACTCGCGCAAAAAACCTCA 4204
Oy 1438 CCGCGAGAGCGGAAGGGGTAAGCGCGCGGGGCTCAAGAGATGCGGGGTCTGATGGC 1497
||||| 4203 CCGCGAAGAGAAAAAATAAATAAAGCAAGAAATCAAAAAATGAAAAATTAATTAAC 4144
Oy 1498 CAGAGTGAAGTGGGTGACCGCGGGGTGCTGAGTGGCTGGAATGCGCGGGGTGGCGGG 1557
||||| 4143 CAATAATTAATTAATTAACGAAATATTAATTAATTAATTAATTAATTAATTAATTAAC 4084
Oy 1558 ACACTTGTCTTGTATGACAGCAGAGTGA--CGGCTCGGGCGGTGAGCAGAGAGCA 1615
||||| 4083 ACACTTGAATTTCTTATTAACAGCAAAATTAACCGAATCTCGAAGCTGACGCAAAAAACA 4024
Oy 1616 GGTGCGGGCGAGTCTGAGTCTGACGCGCCCGGAGCTGGGCGGGCTGCTGAGGGCG 1675
||||| 4023 AATACCGAAGCAATTTCTGACTACGCGCCCGAAGTACGCGCGAGCTACTAATAAACA 3964
Oy 1676 AGCAGC-GGCGCGCGACCGCTGCTGCGCGAGACCGCCCTCCCAACCCCACTCTTG 1734
||||| 3963 AACAGAACCCCGCGAAGCGTACTACGCGAGACCGCCCTCCCAACCCCACTCTTA 3904
Oy 1735 GCGCGC-CGTTCCGGGGGCTGTCTGGCGACCCCGGCTTCTATATACGGGCGCGCG 1793
||||| 3903 AACGCGCGCTTCCGAACGATCTTAACACCCCGACTTCTATTAACGACCGAGCGCG 3844
Db 1794 CCGGCGCGGCCA 1805
||||| 3843 CCGAACCGGCCA 3832
Db 3843 CCGAACCGGCCA 3832

RESULT 7
AF305712 13720 bp DNA linear MAM 05-NOV-2000
LOCUS AF305712 Bos laurus insulin-like growth factor binding protein-3 (IGFBP3)
DEFINITION Bos laurus insulin-like growth factor binding protein-3 (IGFBP3)
ACCESSION AF305712
VERSION AF305712.1 GI:11095302
KEYWORDS
SOURCE
ORGANISM
Bos laurus.
Bos laurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 13720)
REFERENCE
Erondu,N.E., Toland,B., Boes,M., Dake,B., Moser,D.R. and Bar,R.S.
Bovine insulin-like growth factor binding protein-3: organization
of the chromosomal gene and functional analysis of its promoter
Endocrinology 138 (7), 2856-2862 (1997)
JOURNAL MEDLINE
PUBMED 97345736
9202228
2 (bases 1 to 13720)
Erondu,N.E., Toland,B., Boes,M., Dake,B. and Bar,R.S.

```


Db 4265 -----GGCGCGCGCTCCGGGCGCTGCTCCGGCGG-CCCGGCTTCTATATACGGCGG 4318
Oy 1789 CGCGCCCGGGCGCCGCC 1804
Db 4319 CTGCGCCCGGACCGCC 4334

RESULT 8
AC109771/c
LOCUS
DEFINITION
AC109771 107161 bp DNA linear HTG 17-JUL-2002
Rattus norvegicus clone CH230-71E4, *** SEQUENCING IN PROGRESS ***,
63 unordered pieces.
AC109771
AC109771.3 GI:21744354
HTG: HTGS_PHASE1.
Norway rat.
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 107161)
Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,D., Bowie,S., Bieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Clelland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.R.,
Donahewaiter,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homi,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsone,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtenage,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Teltord,B., Thomas,N., Thomas,S.,
Umanil,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlingon,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (17-JUL-2002) Human Genome Sequencing Center, Department

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 13, 2002 this sequence version replaced gi:18860445.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRPC
Center clone name: CH230-71E4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 50607 bases at least Q40
Consensus quality: 54380 bases at least Q30
Consensus quality: 58166 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1105: contig of 1105 bp in length
1205: gap of unknown length
1206: contig of 1036 bp in length
2242: gap of unknown length
2342: gap of unknown length
3377: gap of unknown length
3476: gap of unknown length
5068: contig of 1592 bp in length
5168: gap of unknown length
5169: contig of 1477 bp in length
6645: gap of unknown length
6745: contig of 1013 bp in length
7758: gap of unknown length
7859: contig of 1173 bp in length
9031: gap of unknown length
9032: contig of 1109 bp in length
9132: gap of unknown length
10241: gap of 1627 bp in length
10241: contig of 1627 bp in length
11967: gap of unknown length
11968: gap of unknown length
12067: gap of unknown length
12068: gap of unknown length
13551: contig of 1484 bp in length
13552: gap of unknown length
13651: gap of unknown length
14771: contig of 1119 bp in length
14771: gap of unknown length
14870: gap of unknown length
14871: contig of 1338 bp in length
16208: gap of unknown length
16209: gap of unknown length
16309: contig of 1160 bp in length
17468: gap of unknown length
17469: gap of unknown length
17568: gap of unknown length
18954: contig of 1366 bp in length
18955: gap of unknown length
19055: gap of unknown length
20323: contig of 1269 bp in length
20324: gap of unknown length
20423: gap of unknown length
21701: contig of 1278 bp in length
21801: gap of unknown length
21802: contig of 1262 bp in length
23063: gap of unknown length
23163: gap of unknown length
23164: contig of 1324 bp in length
24487: gap of unknown length
24587: gap of unknown length
24588: gap of unknown length
25769: contig of 1182 bp in length
25869: gap of unknown length
25870: gap of unknown length
27369: contig of 1500 bp in length
27469: gap of unknown length
27470: gap of unknown length
29433: contig of 1964 bp in length
29434: gap of unknown length
29533: gap of unknown length
31050: contig of 1517 bp in length
31051: gap of unknown length

31151 32640: contig of 1490 bp in length
32641 32740: gap of unknown length
32741 34019: contig of 1279 bp in length
34020 34119: gap of unknown length
34120 36105: contig of 1986 bp in length
36106 36205: gap of unknown length
36206 37278: contig of 1073 bp in length
37279 37378: gap of unknown length
37379 39416: contig of 2038 bp in length
39417 41021: contig of 1505 bp in length
41022 41121: gap of unknown length
41122 42627: contig of 1506 bp in length
42628 42727: gap of unknown length
42728 44369: contig of 1642 bp in length
44370 44469: gap of unknown length
44470 45561: contig of 1092 bp in length
45562 45661: gap of unknown length
45662 46976: contig of 1315 bp in length
46977 47076: gap of unknown length
47077 48433: contig of 1267 bp in length
48434 49674: contig of 1231 bp in length
49675 49774: gap of unknown length
49775 51642: contig of 1868 bp in length
51643 51742: gap of unknown length
51743 53097: contig of 1355 bp in length
53098 53197: gap of unknown length
53198 55117: contig of 1920 bp in length
55118 55217: gap of unknown length
55218 57573: contig of 2356 bp in length
57574 57673: gap of unknown length
57674 59663: contig of 1890 bp in length
59664 61822: contig of 2159 bp in length
61823 61922: gap of unknown length
61923 63079: contig of 1157 bp in length
63080 63179: gap of unknown length
63180 64615: contig of 1436 bp in length
64616 67307: contig of 2592 bp in length
67308 67407: gap of unknown length
67408 68773: contig of 1366 bp in length
68774 68873: gap of unknown length
68874 70712: contig of 1839 bp in length
70713 70812: gap of unknown length
70813 73004: contig of 2192 bp in length
73005 73104: gap of unknown length
73105 74453: contig of 1349 bp in length
74454 74553: gap of unknown length
74554 76324: contig of 1771 bp in length
76325 76424: gap of unknown length
76426 77737: contig of 1313 bp in length
77738 77837: gap of unknown length
77839 79745: contig of 1908 bp in length
79746 79845: gap of unknown length
79846 81448: contig of 1603 bp in length
81449 81548: gap of unknown length
81549 83338: contig of 1790 bp in length
83339 83438: gap of unknown length
83439 84877: contig of 1433 bp in length
84878 84977: gap of unknown length
84978 87020: contig of 2043 bp in length

Query Match 4.6%; Score 82.6; DB 2; Length 107161;
Best Local Similarity 60.0%; Pred. No. 1.4e-08;
Matches 219; Conservative 0; Mismatches 124; Indels 22; Gaps 4;

QY 1445 AGCGAAGGCGTAAGGCGGCGGCTCAAGAGATCGGGGTGCTGAGTTGCCAGAGT 1504
DB 67899 AGAGGAGGAGCGGTGAAGGCGGCGGAGCCCAAGTTGGCTTAAGAGTCTTTGGGGAGAGAGA 67840
QY 1505 GACTGGGGTGACCGGGGGGTGCTGAGTGGCTGAGTCCGGGGGTGGCCGGGACACACT- 1563

DB 67839 GAACAGAGCCCCGTGGATGCTGGACTCCAGAGACAGCAAGTAGCCTGGAGCCCTG 67780
QY 1564 TGTCTTCTTGTAGCAAGAGTG--ACGGGCTCGGCGCTGAGACAGAGAGAGGTGCC 1621
DB 67779 CCGTTACCGGGGAGCAAGAAAGTGGCCCTAGTCTGAGAGTCCAGAGAGAGAGTGGC 67720
QY 1622 CGGGCGAGTCAGAGTCAGACAGCCCGGAGCTGGCCCGGCTGTCAGAGCGCAAGCAG 1681
DB 67719 CAGTGAAGACAGCGGAGCGGCGGCGCCGCTGCTGATCCGCTTCGGCCG---GAGGCGGG 67663
QY 1682 GGGCCCGAGCCGCTGCTGGCGGAGCCCGCCCTTCCCAACCCCACTCTCGGCGCC 1741
DB 67662 AGCCGGGAGCGAGCCTGGCTGGCCACCCCGCCAGCCCGC-----C 67619
QY 1742 GTTCGGGGGCTGCTGCGGCACCCCGGCTTATATACGGCGGCGGCGCCGCGCG 1801
DB 67618 GTTCGGGGGCTGCTGCTGCGGCACCCCGGCTTATATATAGGCTGCGGCGCGCACTG 67559
QY 1802 CCCAG 1806
DB 67558 CCCAG 67554

RESULT 9
LOCUS 166494/c 7218 bp DNA linear PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
VERSION 166494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source 1. 7218
location/Qualifiers
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 4.4%; Score 78.6; DB 6; Length 7218;
Best Local Similarity 11.2%; Pred. No. 8.3e-08;
Matches 30; Conservative 159; Mismatches 78; Indels 0; Gaps 0;

QY 256 TGTAGAGCTCAATTATTTATTTATTTAGTACAGAGGAGCAAGTAATGATCTTCCATT 315
DB 1496 TGTAGGATCATCTGATTTACTATCTATGCAAGTAGTTAAAGATAGTAAGTAATTTGGT 1437
QY 316 CTAAAGAGAGATATATATGTAATAAGAGCTGATAGATATGGGGAAGAGTGAGC 375
DB 1436 ACRR 1377
QY 376 AGGGGGAAGAGGAG 435
DB 1376 RRR 1317
QY 436 TGGCGAGAGATGATTTAGAGAGACAGAAATGATGATTAAGATTAACTTCAACCAAG 495
DB 1316 RRR 1257
QY 496 CTTCGTGCTGAGGAGAGATGAGAGAG 522
DB 1256 RRR 1230

RESULT 10
LOCUS ACOT8884 207420 bp DNA linear HTG 15-MAY-2002
DEFINITION Mus musculus chromosome 17 clone RP23-390F1 strain C57BL6/J,

[illegible]

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I17376
Center clone name: 100_D_4

* NOTE: This record contains 85 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
725 824: gap of 100 bp in length
825 1513: contig of 689 bp in length
1514 1613: gap of 100 bp
1614 2330: contig of 717 bp in length
2331 2430: gap of 100 bp
2431 3129: contig of 699 bp in length
3130 3229: gap of 100 bp
3230 3928: contig of 699 bp in length
3929 4028: gap of 100 bp
4029 4754: contig of 726 bp in length
4755 4854: gap of 100 bp
4855 5567: contig of 713 bp in length
5568 5667: gap of 100 bp
5668 6396: contig of 729 bp in length
6397 6496: gap of 100 bp
6497 7224: contig of 728 bp in length
7225 7324: gap of 100 bp
7325 8038: contig of 714 bp in length
8039 8138: gap of 100 bp
8139 8846: contig of 708 bp in length
8847 8946: gap of 100 bp
8947 9683: contig of 737 bp in length
9684 9783: gap of 100 bp
9784 10482: contig of 699 bp in length
10483 10582: gap of 100 bp
10583 11306: contig of 724 bp in length
11307 11406: gap of 100 bp
11407 12125: contig of 719 bp in length
12126 12225: gap of 100 bp
12226 12936: contig of 711 bp in length
12937 13036: gap of 100 bp
13037 13744: contig of 708 bp in length
13745 13844: gap of 100 bp
13845 14553: contig of 709 bp in length
14554 14653: gap of 100 bp
14654 15362: contig of 709 bp in length
15363 15462: gap of 100 bp
15463 16183: contig of 721 bp in length
16184 16283: gap of 100 bp
16284 17005: contig of 722 bp in length
17006 17105: gap of 100 bp
17106 17824: contig of 719 bp in length
17825 17924: gap of 100 bp
17925 18628: contig of 704 bp in length
18629 18728: gap of 100 bp
18729 19440: contig of 712 bp in length
19441 19540: gap of 100 bp
19541 20261: contig of 721 bp in length
20262 20361: gap of 100 bp
20362 21061: contig of 700 bp in length
21062 21161: gap of 100 bp
21162 21876: contig of 715 bp in length
21877 21976: gap of 100 bp
21977 22677: contig of 701 bp in length
22678 22777: gap of 100 bp
22778 23507: contig of 730 bp in length
23508 23607: gap of 100 bp
23608 24317: contig of 710 bp in length
24318 24417: gap of 100 bp
24418 25128: contig of 711 bp in length
25129 25228: gap of 100 bp
25229 25944: contig of 716 bp in length
25945 26044: gap of 100 bp
26045 26772: contig of 728 bp in length
26773 26872: gap of 100 bp
26873 27576: contig of 704 bp in length
27577 27676: gap of 100 bp
27677 28377: contig of 701 bp in length

28378 28477: gap of 100 bp
28478 29181: contig of 704 bp in length
29182 29281: gap of 100 bp
29282 30001: contig of 720 bp in length
30002 30101: gap of 100 bp
30102 30822: contig of 721 bp in length
30823 30922: gap of 100 bp
30923 31646: contig of 724 bp in length
31647 31746: gap of 100 bp
31747 32463: contig of 717 bp in length
32464 32563: gap of 100 bp
32564 33281: contig of 718 bp in length
33282 33381: gap of 100 bp
33382 34085: contig of 704 bp in length
34086 34185: gap of 100 bp
34186 34910: contig of 725 bp in length
34911 35010: gap of 100 bp
35011 35729: contig of 719 bp in length
35730 35829: gap of 100 bp
35830 36547: contig of 718 bp in length
36548 36647: gap of 100 bp
36648 37344: contig of 697 bp in length
37345 37444: gap of 100 bp
37445 38149: contig of 705 bp in length
38150 38249: gap of 100 bp
38250 38956: contig of 707 bp in length
38957 39056: gap of 100 bp
39057 39766: contig of 710 bp in length
39767 39866: gap of 100 bp
39867 40563: contig of 697 bp in length
40564 40663: gap of 100 bp
40664 41387: contig of 724 bp in length
41388 41487: gap of 100 bp
41488 42206: contig of 719 bp in length
42207 42306: gap of 100 bp
42307 42982: contig of 676 bp in length
42983 43082: gap of 100 bp
43083 43787: contig of 705 bp in length
43788 43887: gap of 100 bp
43888 44597: contig of 710 bp in length
44598 44697: gap of 100 bp
44698 45423: contig of 726 bp in length
45424 45523: gap of 100 bp
45524 46238: contig of 715 bp in length
46239 46338: gap of 100 bp
46339 47056: contig of 718 bp in length
47057 47156: gap of 100 bp
47157 47868: contig of 712 bp in length
47869 47968: gap of 100 bp
47969 48679: contig of 711 bp in length
48680 48779: gap of 100 bp
48780 49483: contig of 704 bp in length
49484 49583: gap of 100 bp
49584 50268: contig of 685 bp in length
50269 50368: gap of 100 bp
50369 51092: contig of 724 bp in length
51093 51192: gap of 100 bp
51193 51909: contig of 717 bp in length
51910 52009: gap of 100 bp
52010 52713: contig of 704 bp in length
52714 52813: gap of 100 bp
52814 53502: contig of 689 bp in length
53503 53602: gap of 100 bp
53603 54330: contig of 728 bp in length
54331 54430: gap of 100 bp
54431 55141: contig of 711 bp in length
55142 55241: gap of 100 bp
55242 55961: contig of 720 bp in length
55962 56061: gap of 100 bp

Query Match 4.1%; Score 74.2; DB 2; Length 68866;
Best Local Similarity 41.4%; Pred. No. 1.3e-06;
Matches 164; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthett, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homel, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Lounsbury, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mayhew, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, N., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, E., Nguyen, N., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, D., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G.,
 Scherer, S., Scott, G., Shen, H., Shooshair, N., Slisdon, I.,
 Sodegriem, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinton, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

Unpublished
 Direct Submission
 2 (bases 1 to 187521)
 Worley, K.C.
 Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 187521)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 30, 2002 this sequence version replaced gi:21629151.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: G2E2
 Center clone name: CH230-204D14

----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 131060 bases at least Q40
 Consensus quality: 138483 bases at least Q30
 Consensus quality: 143840 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 73 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

* be preserved.
 1
 1213
 1313
 1313
 2942
 2942
 3042
 4406
 4406
 4506
 4506
 4507
 6177
 6177
 6277
 7554
 7554
 9023
 9123
 10550
 10650
 11655
 11655
 11765
 11765
 12783
 12883
 13945
 14046
 14046
 15825
 15825
 15925
 17209
 17309
 18987
 18987
 20250
 20350
 22358
 22358
 22458
 23772
 23772
 25528
 25528
 26804
 26804
 28445
 28445
 30094
 30094
 30194
 32189
 32189
 32290
 32290
 33290
 33290
 33305
 33305
 33406
 33406
 34911
 34911
 35011
 35011
 36646
 36646
 38285
 38285
 38285
 40884
 40884
 40885
 40885
 42961
 42961
 43061
 43061
 44757
 44757
 44858
 44858
 46199
 46199
 46299
 46299
 48724
 48724
 48725
 50474
 50474
 50574
 52657
 52657
 52757
 54755
 54755
 54855
 54855
 56446
 56446
 58340
 58340
 58441
 58441
 59626
 59626

1212: contig of 1212 bp in length
 1312: gap of unknown length
 1313: contig of 1629 bp in length
 2941: gap of unknown length
 2942: gap of unknown length
 3041: gap of unknown length
 4405: contig of 1365 bp in length
 4406: gap of unknown length
 4505: gap of unknown length
 4506: contig of 1670 bp in length
 6176: gap of unknown length
 6177: contig of 1377 bp in length
 7553: gap of unknown length
 7554: contig of 1265 bp in length
 9022: gap of unknown length
 9122: gap of unknown length
 10549: contig of 1427 bp in length
 10649: gap of unknown length
 11654: contig of 1015 bp in length
 11764: gap of unknown length
 12782: contig of 1018 bp in length
 12882: gap of unknown length
 13945: contig of 1063 bp in length
 14045: gap of unknown length
 15824: contig of 1779 bp in length
 15924: gap of unknown length
 17208: contig of 1284 bp in length
 17308: gap of unknown length
 18986: contig of 1678 bp in length
 19086: gap of unknown length
 20249: contig of 1163 bp in length
 20349: gap of unknown length
 22357: contig of 2008 bp in length
 22457: gap of unknown length
 23671: contig of 1214 bp in length
 23771: gap of unknown length
 25527: contig of 1756 bp in length
 25627: gap of unknown length
 26803: contig of 1176 bp in length
 26903: gap of unknown length
 28445: contig of 1542 bp in length
 28445: gap of unknown length
 30093: contig of 1548 bp in length
 30193: gap of unknown length
 32189: contig of 1996 bp in length
 32289: gap of unknown length
 33305: contig of 1016 bp in length
 33405: gap of unknown length
 34910: contig of 1505 bp in length
 35010: gap of unknown length
 36645: contig of 1635 bp in length
 36745: gap of unknown length
 38285: contig of 1540 bp in length
 38285: gap of unknown length
 40884: contig of 2499 bp in length
 40884: gap of unknown length
 42961: contig of 1977 bp in length
 43061: gap of unknown length
 44757: contig of 1696 bp in length
 44857: gap of unknown length
 46198: contig of 1341 bp in length
 46298: gap of unknown length
 48724: contig of 2326 bp in length
 48724: gap of unknown length
 50473: contig of 1749 bp in length
 50573: gap of unknown length
 52656: contig of 2083 bp in length
 52756: gap of unknown length
 54754: contig of 1998 bp in length
 54854: gap of unknown length
 56445: contig of 1591 bp in length
 56445: gap of unknown length
 58340: contig of 1795 bp in length
 58440: gap of unknown length
 59625: contig of 1165 bp in length
 59725: gap of unknown length

```

* 59726 62021: contig of 2296 bp in length
* 62022 62121: gap of unknown length
* 62122 63651: contig of 1530 bp in length
* 63652 63751: gap of unknown length
* 63752 66236: contig of 2485 bp in length
* 66237 66336: gap of unknown length
* 66337 68540: contig of 2204 bp in length
* 68541 68641: gap of unknown length
* 68641 70800: contig of 2160 bp in length
* 70801 73422: contig of 2522 bp in length
* 73423 73522: gap of unknown length
* 73523 76025: contig of 2502 bp in length
* 76025 76125: gap of unknown length
* 76125 77614: contig of 1389 bp in length
* 77614 79925: contig of 2312 bp in length
* 79926 80025: gap of unknown length
* 80026 82184: contig of 2159 bp in length
* 82185 82284: gap of unknown length
* 82285 83574: contig of 1290 bp in length
* 83575 85573: contig of 1899 bp in length
* 85574 85674: gap of unknown length
* 85674 88158: contig of 2485 bp in length
* 88159 88259: gap of unknown length
* 88259 91411: contig of 3153 bp in length
* 91412 91511: gap of unknown length
* 91512 94235: contig of 2724 bp in length
* 94236 94335: gap of unknown length
* 94336 96689: contig of 2354 bp in length
* 96690 96789: gap of unknown length
* 96790 100020: contig of 3231 bp in length
* 100021 100121: gap of unknown length
* 100121 102772: contig of 2652 bp in length

Query Match      3.8%; Score 68.2; DB 2; Length 187521;
Best Local Similarity 47.6%; Pred. No. 4.2e-05;
Matches 166; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```

```

AL591129/c 213985 bp DNA linear ROD 29-JUN-2002
LOCUS Mouse DNA sequence from clone Rp23-147P4 on chromosome 11, complete
DEFINITION
ACCESSION AL591129
VERSION AL591129.24 GI:21665918
KEYWORDS HTG.

```

```

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 213985)
JOURNAL Direct Submission
COMMENT Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humqueres@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On Jul 2, 2002 this sequence version replaced gi:21653323.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humqueres@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep Rp23-147P4 is from the Rp23-Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6.

```

FEATURES
    source
        1..213985
            /location/Qualifiers
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /chromosome="11"
                /clone="Rp23-147P4"
                /clone_1lb="Rp23-23"

```

```

BASE COUNT      62282 a 47467 c 48155 g 56081 t
ORIGIN

```

```

Query Match      3.8%; Score 68.2; DB 10; Length 213985;
Best Local Similarity 56.4%; Pred. No. 4.3e-05;
Matches 127; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```

```

QY 256 TGTAGAGGTCATTTATTTATTTAGTACAAAGGGAGAGAAATGATTTTCATT 315
|||||
Db 165045 TGTAGAGGTCGTGTGATCTTGAGTTTGTCGACATCCAAATCTAGTACCTGT 164986
|||||
QY 316 CTAAGAGAGATATATATATTAAGAGAGCTGTATAGATATGGGGAGAGGTGAC 375
|||||
Db 164985 CTCATAAAGAGAGATATATATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 164926
|||||
QY 376 AGGGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 435
|||||
Db 164925 AGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 164866
|||||
QY 436 TGGGCGAGATGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
|||||
Db 164865 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 164821
|||||

```

```

AL806529/c 214511 bp DNA linear HTG 17-AUG-2002
LOCUS

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:54:28 ; Search time 408 Seconds
(without alignments)
9968.404 Million cell updates/sec

Title: US-09-865-879-1

Perfect score: 1806
Sequence: 1 gggagctgttcttcttccct.....ggcgccggcgccgcccag 1806

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806	100.0	1806	24	AAD24912
2	1795	99.4	10884	24	ABN95848
3	1141.4	63.2	8700	24	ABL33002
4	1021.2	56.5	8700	24	ABL33003
5	117	6.5	1007	24	ABO45156
6	117	6.5	1007	24	ABO45157
7	84.4	4.7	1007	24	ABO45154
8	84.4	4.7	1007	24	ABO45155
9	60.8	3.4	6027	24	ABN80173

10	59	3.3	395	22	AAF64375	Novel human polynu
11	59	3.3	16200	24	AAD36438	Mouse L66 genomic
12	58.4	3.2	170	21	ABN81171	Shrimp polynucleot
13	57.8	3.2	502	21	ABN81079	Shrimp polynucleot
14	57.4	3.2	396	22	AAF64796	Novel human polynu
15	57	3.2	420	22	AAF67104	Novel human polynu
16	56.6	3.1	307	24	ABL86438	Human ovarian canc
17	56.6	3.1	5191	21	AAZ98617	RPPI-MSA genomic n
18	56.6	3.1	10301	22	AAK84949	Human immune/Haema
19	56.2	3.1	329	17	AAT12183	Partial pUG4-5-CDK
20	56.2	3.1	384	22	AAF64952	Novel human polynu
21	56.2	3.1	421	22	AAF66755	Novel human polynu
22	56.2	3.1	36301	20	AAZ23892	Murine LOB0 genom
23	56.2	3.1	38886	20	AAZ23897	Murine LOB0 homolo
24	56.2	3.1	611590	21	AAF22303	Arabidopsis thalia
25	56	3.1	395	22	AAF66094	Novel human polynu
26	56	3.1	611590	21	AAF22303	Arabidopsis thalia
27	55.6	3.1	298	24	ABL86000	Human ovarian can
28	55.6	3.1	397	22	AAF64573	Novel human polynu
29	55.6	3.1	863	20	AAZ17377	Human gene express
30	55.6	3.1	3586	24	AAZ63366	Chemically pretrea
31	55.4	3.1	398	22	AAF64279	Novel human polynu
32	55.4	3.1	400	22	AAF64260	Novel human polynu
33	55.4	3.1	422	22	AAF66607	Novel human polynu
34	55.4	3.1	424	22	AAF67052	Novel human polynu
35	55.2	3.1	390	22	AAF64933	Novel human polynu
36	55.2	3.1	399	22	AAF66668	Novel human polynu
37	55.2	3.1	405	22	AAF64515	Novel human polynu
38	55.2	3.1	4752	15	AAO58032	Tyrosinase promote
39	55	3.0	5306	24	ABL32510	Human immune syste
40	54.8	3.0	190	16	AAZ45061	Ts2 gene microsat
41	54.6	3.0	300	20	AAZ13219	Human gene express
42	54.6	3.0	382	22	AAF65098	Novel human polynu
43	54.6	3.0	700	11	AAO05263	Sequence of neuron
44	54.6	3.0	700	15	AAO57505	Rat GAP-43 promote
45	54.6	3.0	700	15	AAO57506	Rat GAP-43 promote

ALIGNMENTS

RESULT 1	
AAD24912	
ID	AAD24912 standard; DNA; 1806 BP.
XX	
XX	AAD24912;
XX	
DT	12-MAR-2002 (first entry)
XX	
XX	
DE	Human insulin-like growth factor binding protein-3 promoter DNA.
XX	
KW	Human: growth inhibitory gene; retinoid; retinoid acid response element;
KW	RARE site; therapy; promyelocytic leukaemia; cancer chemoprevention;
KW	cytostatic; insulin-like growth factor binding protein; IGFBP-3 promoter;
KW	ds.
OS	
OS	Homo sapiens.
XX	
PN	MO200192578-A2.
PD	
PD	06-DEC-2001.
XX	
PF	25-MAY-2001; 2001WO-US17161.
XX	
PR	26-MAY-2000; 2000US-207535P.
XX	
PA	(UNIT) UNITV ILLINOIS FOUNO.
XX	
PI	Roninson IB, Dokmanovic M, Chang B;
XX	
DR	WPI; 2002-075474/10.
XX	
PT	Expression construct encoding cellular genes, under control of a

DB 1801 GCCCAG 1806

RESULT 2
ABN95848
ID ABN95848 standard; DNA: 10884 BP.
XX
AC ABN95848;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2346 used to diagnose liver cancer.
XX
KW Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-De-Silva S, Vockley JG;
XX
DR WPI: 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 2346; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 10884 BP; 2796 A; 2578 C; 2737 G; 2773 T; 0 other;

Query Match 99.4%; Score 1795; DB 24; Length 10884;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1806; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGGGATTCGTTTGGTTCCCTTCATTTTCCAAATGAAATCAGAGATCCTGTTGGGTGT 60
DB 101 GGGGATTCGTTTGGTTCCCTTCATTTTCCAAATGAAATCAGAGATCCTGTTGGGTGT 160

QY 61 CAACGAGATTAAGAGAGAGTGTATACAGAGAAAG- AAACGACAGGAGCATATATG 119
DB 161 CAACGAGATTAAGAGAGAGTGTATACAGAGAAAG- AAACGACAGGAGCATATATG 220

QY 120 CACGGTTTCCTTAACAGGTGTGAGTGTAGCCACAGCCTGTGAGCACTGTGGGAGAGAC 179
DB 221 CACGGTTTCCTTAACAGGTGTGAGTGTAGCCACAGCCTGTGAGCACTGTGGGAGAGAGAC 280

QY 180 TCATAGAAAATAGCGGTGCTGGCCCTTCGTACACCCCGGGCCCTCCATTTGTTCT 239
DB 281 TCATAGAAAATAGCGGTGCTGGCCCTTCGTACACCCCGGGCCCTCCATTTGTTCT 340

QY 240 TTGGTCTCTTTTATTTGTAAGAGTCCAAATTTATTTATTTATTAAGAGGAAAGCA 299
DB 341 TTGGTCTCTTTTATTTGTAAGAGTCCAAATTTATTTATTTATTTATTAAGAGGAAAGCA 400

QY 300 AATTGATCTTTCATCTTAAGAGAGTATATATGATATTAAGAGAGCGTATATAT 359
DB 401 AATTGATCTTTCATCTTAAAGAGAGTATATATGATATTAAGAGAGCGTATATAT 460

QY 360 GGGGAGAGAGGTGAGACAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
DB 461 GGGGAGAGAGGTGAGACAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520

QY 420 GGACAAAGAGAGACACTGGCGGAGATCGATTAGAGAGACAGAAATGATGAATGAAGA 479
DB 521 GGACAAAGAGAGACACTGGCGGAGATCGATTAGAGAGAGAGAGAGAGAGAGAGAG 580

QY 480 TTACCTTCACCCCAAGGCTTCGCTGGAGAGGGAATGAGAGAGAGAGAGAGAGAGAG 539
DB 581 TTACCTTCACCCCAAGGCTTCGCTGGAGAGGGAATGAGAGAGAGAGAGAGAGAGAG 640

QY 540 CTACTCCAAACTGCAAAAGGCTCTTCAAGTCACTATCCACTCCCTAAGGCAAGGCTCC 599
DB 641 CTACTCCAAACTGCAAAAGGCTCTTCAAGTCACTATCCACTCCCTAAGGCAAGGCTCC 700

QY 600 AATTTCACAGCGTTGAGAGAAAGTCTCCGCGAGAGTCTACCGCTTCCACTCCAC 659
DB 701 AATTTCACAGCGTTGAGAGAAAGTCTCCGCGAGAGTCTACCGCTTCCACTCCAC 760

QY 660 CCCCAAAACTCTTGGAAAAGTGCCTGAAAATTTAAATCCATTCATTCCTGAGAC 719
DB 761 CCCCAAAACTCTTGGAAAAGTGCCTGAAAATTTAAATCCATTCATTCCTGAGAC 820

QY 720 ACCAGGCTCTCTGTTGGTCCACCGAAGAGAGGGGTGCGCAGACAAAACGTGAAAACCTCG 779
DB 821 ACCAGGCTCTCTGTTGGTCCACCGAAGAGAGGGGTGCGCAGACAAAACGTGAAAACCTCG 880

QY 780 AGTGCAGAGAGAGCCGACAGAGAGTATACACGACCTCAGGCGCAATTTGGCCCGAACT 839
DB 881 AGTGCAGAGAGAGCCGACAGAGAGTATACACGACCTCAGGCGCAATTTGGCCCGAACT 940

QY 840 TTACTGAAAAGTGTATAGATTGAGAGATGAGATAGATCCCAAGCATGAGAGATACAG 899
DB 941 TTACTGAAAAGTGTATAGATTGAGAGATGAGATAGATCCCAAGCATGAGAGATACAG 1000

QY 900 TAATACGAAGTGCCTTCAAAAATGACAAATGAAAATGCGATTAAAGACTATTTGGT 959
DB 1001 TAATACGAAGTGCCTTCAAAAATGACAAATGAAAATGCGATTAAAGACTATTTGGT 1060

QY 960 TAATACGTTTACAGAGTGCCTTCAAGTTATTTGTTTATTTTGTGCTGGGTGTA 1019
DB 1061 TAATACGTTTACAGAGTGCCTTCAAGTTATTTGTTTATTTTGTGCTGGGTGTA 1120

QY 1020 ACTCCATTTGAAAACATTAATCAGAGAGATACCAAGAGAGAGAGAGAGAGAGAGAGAG 1079
DB 1121 ACTCCATTTGAAAACATTAATCAGAGAGATACCAAGAGAGAGAGAGAGAGAGAGAGAG 1180

QY 1080 AATATTTGAAAAGCCCTCCTTAAGAGACATTCGCTGCGGCTCCACTTTAATTTGGG 1139
DB 1181 AATATTTGAAAAGCCCTCCTTAAGAGACATTCGCTGCGGCTCCACTTTAATTTGGG 1240

QY 1140 ACTTGGGTGTGACAAACAGTGAAGTCTTCTTGGCTTGAAGAGTAAAGCTGGAAGGCG 1199
DB 1241 ACTTGGGTGTGACAAACAGTGAAGTCTTCTTGGCTTGAAGAGTAAAGCTGGAAGGCG 1300

QY 1200 AAGGCCCCGGGAGATCTTGAATGAGTATTTGTTGGGCCCCGAGGATTAACAGCCGAG 1259
DB 1301 AAGGCCCCGGGAGATCTTGAATGAGTATTTGTTGGGCCCCGAGGATTAACAGCCGAG 1360

QY 1260 CCGGTGTAATTAACCCGAGTGCCTGGCTCCTGAGACCCAAATGTAAGTCAGAAA 1319
 DB 1361 CCGGTGTAATTAACCCGAGTGCCTGGCTCCTGAGACCCAAATGTAAGTCAGAAA 1420
 QY 1320 TGGCCAGAGCTTCCGCTGCAACGGAATTAATTTAGAAAGCTCCACGAGGTACAC 1379
 DB 1421 TGTCCAGAGCTTCCGCTGCAACGGAATTAATTTAGAAAGCTCCACGAGGTACAC 1480
 QY 1380 GAATGGGAGCGCTGATGACAGTTCCTCCGACACCGGCTCCGCGACGAGGAGACCTCACC 1439
 DB 1481 GAATGGGAGCGCTGATGACAGTTCCTCCGACACCGGCTCCGCGACGAGGAGACCTCACC 1540
 QY 1440 CCGAGAGCGGAGGAGGATTAAGGCGCGGGCTCAAGAGATCGGGGCTGCTAGTTGGCCA 1499
 DB 1541 CCGAGAGCGGAGGAGGATTAAGGCGCGGGCTCAAGAGATCGGGGCTGCTAGTTGGCCA 1600
 QY 1500 GGAGTACTGAGGAGTGAACCGGGGCTGAGAGTGGAGTGCCTGGGAGTGGCGGGGAC 1559
 DB 1601 GGAGTACTGAGGAGTGAACCGGGGCTGAGAGTGGAGTGCCTGGGAGTGGCGGGGAC 1660
 QY 1560 ACCCTGGTCTTGTAGACAGCAAGGTGACGGGCTCCGGGCTGAGCAGAGAGAGGTG 1619
 DB 1661 ACCCTGGTCTTGTAGACAGCAAGGTGACGGGCTCCGGGCTGAGCAGAGAGAGGTG 1720
 QY 1620 CCGGAGGAGTCTGAGCTGACAGCCCGGAGCTCGGCTCGCTCAGGGGGAAGCA 1679
 DB 1721 CCGGAGGAGTCTGAGCTGACAGCCCGGAGCTCGGCTCGCTCAGGGGGAAGCA 1780
 QY 1680 CCGGAGGAGTCTGAGCTGACAGCCCGGAGCTCGGCTCGCTCAGGGGGAAGCA 1739
 DB 1781 CCGGAGGAGTCTGAGCTGACAGCCCGGAGCTCGGCTCGCTCAGGGGGAAGCA 1840
 QY 1740 GCGTTCCGGGCGCTGCTGAGGCGACCCCGGCTCTATATATAGGCGCGCGCGGCG 1799
 DB 1841 GCGTTCCGGGCGCTGCTGAGGCGACCCCGGCTCTATATATAGGCGCGCGCGGCG 1900
 QY 1800 CCGCCAG 1806
 DB 1901 CCGCCAG 1907
 RESULT 3
 ABL33002
 ID ABL33002 standard; DNA: 8700 BP.
 AC ABL33002;
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 975.
 XX
 KW Human; immune system disease; cytosine methylation; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001MO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-103529.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1; SEQ ID NO 975; 32bp + Sequence Listing; German.
 CC
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 8700 BP; 2247 A; 281 C; 2233 G; 3939 T; 0 other;

Query Match 63.2%; Score 1141.4; DB 24; Length 8700;
 Best Local Similarity 80.0%; Pred. No. 1.2e-294;
 Matches 1454; Conservative 0; Mismatches 351; Indels 12; Gaps 9;

QY 1 GGGGATTCGTTTGTCTTCAATTTCCAAATGAAATCAGAGATCCTGTTGGGTG 60
 DB 3055 GGGGATTCGTTTGTCTTCAATTTCCAAATGAAATGAGATTTGTTGGGTG 3114
 QY 61 CAACGCAATCTGAAAGAGGTGATACAGAGAAAGG-AACGCAAGCAGATATATG 119
 DB 3115 TAAAGTATATTAAGAGAGGTGATACAGAGAAAGGAAATGTAAGCAGCATATATG 3174
 QY 120 CACGTTTCTGTAACAAAGTGTAGTCCACAGCCTGAGCAGTGTGAGAGAGAGC 179
 DB 3175 TACGTTTCTGTAACAAAGTGTAGTCCACAGCCTGAGCAGTGTGAGAGAGAGC 3234
 QY 180 TCATTAAGAAATGACGGTGTGGCTTGTCAACCCCGGGGCTCCATTTGTTGTCT 239
 DB 3235 TTATTAAGAAATGACGGTGTGGCTTGTCAACCCCGGGGCTCCATTTGTTGTCT 3294
 QY 240 TTGCTCTCTTTATTTGATGAGTCCAAATTTATTTATTTAGTACAGAGAGAGCA 299
 DB 3295 TTGCTCTCTTTATTTGATGAGTCCAAATTTATTTATTTATTTAGTACAGAGAGCA 3354
 QY 300 AATTGATCTTCCATCTTAAGAGAGATATATATATATTAAGAGAGAGTATATAT 359
 DB 3355 AATTGATCTTCCATCTTAAGAGAGATATATATATATTAAGAGAGAGTATATAT 3414
 QY 360 GGGGAGAGAGTGTGACAGGGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
 DB 3415 GGGGAGAGAGTGTGACAGGGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3474
 QY 420 GGACAGAGAGAGACATCGGCGAGAGATGATTAAGAGAGAGAGAGAGAGAGAGAG 479
 DB 3475 GGATTAAGAGAGAGATGATGAGAGAGATGATTAAGAGAGAGAGAGAGAGAGAG 3534
 QY 480 TTAATCTCACCAAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
 DB 3535 TTAATCTCACCAAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3593
 QY 540 CTACTCCAAATGCAAGGCTCTCAATGACATATCCATCTTAAGAGAGAGAGAG 599
 DB 3594 TTAATCTCACCAAGGCTCTCAATGACATATATATATTTTAAGAGAGAGAGAG 3653
 QY 600 AATTTCACAGAGTTCAGAGAGATCTCCCGGAGAGTTCACCGCTCCACTCAC 659
 DB 3654 AATTTCACAGAGTTCAGAGAGATCTCCCGGAGAGTTCACCGCTCCACTCAC 3713
 QY 660 CCCACAACTCTTGGAAAGTGCCTGAAATTAATTCGCAATCAACCTCGAGAGC 719
 DB 3714 TTTTATTAATTTTGGAAAGTGCCTGAAATTAATTTATTTATTTATTTGGAGT 3773

QY	720	ACGAGCCGCTCTGTGGTACCGAA - GGGAGGGGGCCGACGAAATCGTAAGAAATC	778
Db	3774	ATTAGCGTTTTTGTGTGTTATCGAAGGGAGGGGGTCCGTAGATTAATAATTGAGAAATTC	3833
QY	779	GAGTGCAGAGAAAGCCGACAGGAGTTTACAGCGACCTTCAGCCGCAATTGCG - CCCCANA	837
Db	3834	GAGTGTTAGAGAAAGTCGATNAGGAGTTATAGCATTTTAGCGGTAAATTGCGTTTTCGAA	3893
QY	838	CTTTTACTGAAAAGTGTGTAGATTGCAAGATTAAGTAAATCCACGCATCGAGAATAC	897
Db	3894	TTTTATTGAAAAGGTGTATGATTTGTAGATTAAGTAAATTTTAAACGTATCGAAATAT	3953
QY	898	AGTAAATCGAAGTCGCTTCAAAAATGCAATGAATAATTCCTTTATAAGACATATTG	957
Db	3954	AGTAAATCGAAGTCGTTTTTAAAAAATGATAAAGAAATGTTTTATTAAGGATTAATTG	4013
QY	958	GTTAATTACGTTTACAGCAGTGCACGATTAATTGTCCTTAATATCTTTTGTGTGGGAGT	1017
Db	4014	GTTAATTACGTTTATAGTAGGTTTAGTTATATGTTTTATTTTTTTGTGTGGGAGT	4073
QY	1018	AAACTCATTTGAAAACATATATCAGGAGATATCCAGACACAAGAAACAGTTGCAAT	1077
Db	4074	AAATTTTATTTGAAAATATATATTAGGAGAAATTTAAATTAAGAAATATGTTGTAAT	4133
QY	1078	TAAATATTTTGGAAAAGCCCGCTTAAGG - GCATCGGTGCCGCTGCACCTTAATG	1138
Db	4134	TAAATATTTTGGAAAAGTTTGTTTTATTAAGACGTAATTCGTTTTCGTTATTTTAAATG	4193
QY	1137	GGGACTTGCAGTGTAGCAACACGTGAGAGCTTCCTTGCCTTAGAAGTAAGCTGGAAG	1196
Db	4194	GAGATTTGCGGTGTAGTAATACGTGAGAGTTTTTTTGCCTTAGAAGTAAGTTTGGAAAG	4253
QY	1197	GCG - - - AAGCCCCGGGGCATCTTCAGATTCGTATTGTGGGCCCTGGGGATATAAACA	1253
Db	4254	GGCGGAAGTTTGGGGCATTTTATTTAGATCGAATTTGGGTTTTTGGGGAATATAATA	4313
QY	1254	GCCGAGGGGTGTAAATTAATTAACCCCGCAGTGCCTTGGCTCCTGAGACCCAATGTAAT	1313
Db	4314	GTTTATGGGGGTGTAAATTAATTTCTATGTTTGTGTTTTTGTAGATTAATTAATGTAAT	4373
QY	1314	CAGAAATGTCCCAAGACTTGGCTGCGCAACGGAATTAATTTTAGAAGCTCCACAGAGT	1373
Db	4374	TAGAAATGTTTAAAGTTTCGTTTGTGTTTAAACGGAATTTAAATTTTAGAAAGTTTACGAGCT	4433
QY	1374	ACACAGCAATGCGAGGCGCTGTATGCGCAGTTTCCCGACACGCGCTCGCGAGGAGAC	1433
Db	4434	ATATACGAATGCGAGAGCGTGTATGTAGTTTTTTCGATATCGTTTCGTCGTAAGGAGAT	4493
QY	1434	CTCAGCCCGGAGACGGAAAGGGGTAAAGGGGGGGGGGTCAAGAGATCGGGGCTCTAGT	1493
Db	4494	TTTTATTTTCGAGACCGGAAGGGGTAAAGGGGCGCGGGGTTAAAGAGATCGGGGGTGTGAGT	4553
QY	1494	TGGCAGCAGCTGACTGGGGGAGACCGGGGGGTCGAGAGTGGCGCTGAAGTGCGGGGTGGCC	1553
Db	4554	TGGTTAGGAATGTATTGGGGTGATCGGGGGGTGTGAAGTGTGTTGAAGTGTGCGGGTGCATC	4613
QY	1554	GGGCACACTTGGTCTTGTAGAAGCAAGATGA - - CGGGCTCCGGGGGTGAGCAGCAGG	1611
Db	4614	GGGTAATTTTGGTTTTTGTAGACGATTAAGTGATTCGGGTTTTCGGGGGTGTGTACGAGG	4673
QY	1612	AGCAGGTGCCCCGGCGAGTCTGAGCTGCACGCCCCGAGTGCGCCGCCGCTGTCTAGG	1671
Db	4674	AGTAAAGTGTTCGGGCGAGTTTTCAGATTTTCAGATTTTTCAGATTCGCGTTTCGTTTAGG	4733
QY	1672	GCGAAGACGAGG - CCGCGACCGCGTCGCGGCGACGACCGGCCCCCTCCAAACCCCACT	1730
Db	4734	GCGAAGTACGGGTTTTGTAATGTGTGTGTGCGTCGATTCGTTTTTTTAAATTTTAAAT	4793
QY	1731	CTTGCGGGGCG - CTTTCCGGGGCTGTCTCGGGCCACCCCGCTCTATATAGGGCCGCG	1789
Db	4794	TTTTGGGGCGCTGTTTTCGGGCGCTGTTTTGGGTATTTTCGTTTTTATATACGGTCCGC	4853
QY	1790	GCGCCCGGGCGCCGCAAG 1806	

[illegible]

```

QY 184 AAAAAATGACGGTCTGGGCTTGCACCCGGGCGCTCCATGTGTTCTGCTTGG 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5462 AAAAAAATAAGATGATTAACCTGTCACCCGAAACACATTAATCTTATCTTAA 5403
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 244 TCACTTTTATTTAGAGGCGCAATTTATTTATTTATTTAGTACAGAGGAGCAAAAT 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5402 TCTCTTTTATTTATTAATAATCCATTTATTTATTTATTTATTAACAAAAAAGCAAAAT 5343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 304 GATCTTTTCATTTGAAAGAGAGTATATGTATTAAGAGAGAGCTGTATAGATGAGG 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5342 AATCTTTTCATTTGAAAGAGAGTATATGTATTAAGAGAGAGCTGTATAGATGAGG 5283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 364 GAAGAGTGCAGAGGGGAAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5282 AAAAAAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 424 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5222 AAAAAAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 CTTCACCCAGAGCTTCTGCTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5162 CTTCACCCAGAGCTTCTGCTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 544 TCCAAAGTGCAGAGAGGCTCTTCAAGTACGATTCACCTCTTAAAGCAGAGCTCAAT 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5103 TCCAAAGTGCAGAGAGGCTCTTCAAGTACGATTCACCTCTTAAAGCAGAGCTCAAT 5044
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 604 TCAACAGGCTTCAAGAGAGTCTCTCCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5043 TCAACAGGCTTCAAGAGAGTCTCTCCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 4984
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 664 ACAAACTCTTGGAAAGTGCCTTGAAGAAATTAATCTCAATCAATCTCTGAGCAGCA 723
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4983 ACAAACTCTTGGAAAGTGCCTTGAAGAAATTAATCTCAATCAATCTCTGAGCAGCA 4924
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 724 GGTCTCTGTGTGTGTCACCG-AGAGAGGGGTGCGCAGAGCAAAATGAGAAATCTGAT 782
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4923 AGTCTCTGTATTAATCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 783 GCGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4863 ACCAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4804
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 842 ACTGAAAGTGTATGATTCAGAGAGATAGCTAGATCCCAAGCATCGAGATACAGTA 901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4803 ACTGAAAGTGTATGATTCAGAGAGATAGCTAGATCCCAAGCATCGAGATACAGTA 4744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 902 ATACGAGAGAGCGCTTCAAAAAATGACAAATGAGAAATTCCTATTAAGAGATTTGGTTA 961
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4743 ATACGAGAGAGCGCTTCAAAAAATGACAAATGAGAAATTCCTATTAAGAGATTTGGTTA 4684
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 962 ATTACGTTTCAAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4683 ATTACGTTTCAAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4624
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1022 TCCATTGGAACATATATGAGAGAGATACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1081
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4623 TCCATTGGAACATATATGAGAGAGATACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 4564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1082 ATATTGAAAAAGCGCTGCTTAAGAG-GCATTTGCTTGGCGGTGCTTAAATTTGGGGA 1140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4563 ATATTGAAAAAGCGCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4504
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1141 CTTCGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4503 CTTCGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1198 CGAAGGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4443 GAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1258 AGCGGCTGTAATTTAAACCCCGAGTGGCTTGGCTCCCTGAGAGCCAAATGTAAGTGA 1317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 4383 AAGCAATATTAATTAACCCCGCAATACCTTAATCCCTAAAAACCAATTAATCAAA 4324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1318 AATGCCAAGACTTTCGCTCCCAAGCAATTAATTTTGAAGAGTCCAGAGATACAG 1377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4323 AATATCCCAAGACTTTCGCTCCCAAGCAATTAATTTTGAAGAGTCCAGAGATACAG 4264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1378 ACGAATGCGAGCGCTGTATGCAATTTCCCGACACCGGCTGCGCAGAGAGAGAGAG 1437
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4263 ACGAATGCGAGCGCTGTATGCAATTTCCCGACACCGGCTGCGCAGAGAGAGAGAG 4204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1438 CCGCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4203 CCGCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1498 CAGAGAGTACTGGGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4143 CAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4084
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1558 ACACCTTGTGTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4083 ACACCTTGTGTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4024
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1616 GGTGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4023 AATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1676 AGCAGG-GGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3963 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3904
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1735 GCGCGG-CTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1793
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3903 AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3844
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1794 CCGGCGGCGGCA 1805
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3843 CCGAAGCGGCGCA 3832
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
ABQ45156/c
ID ABQ45156 standard; DNM: 1007 BP.
XX
AC ABQ45156;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 31747.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX drug; side effect; cancer; central nervous system; cardiovascular;
XX gastrointestinal; respiratory system; single nucleotide polymorphism;
XX SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PE 01-SEP-2001; 2001MO-EP10074.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIC-) EPIDENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful

```

Pr for diagnosis and prognosis, comprises selective hybridization of
Pt amplicons from chemically treated DNA -
Pn
XX
PS Claim 12; 56bp + Sequence Listing; 56bp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

OS Homo sapiens.
 XX WO200218632-A2.
 PN
 XX
 PD 07-MAR-2002.
 XX
 XX 01-SEP-2001; 2001WO-EP10074.
 PF
 XX 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K, Guetig D:
 PI
 DR WPI; 2002-371829/40.
 XX
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 CC
 XX
 XX Sequence 1007 BP; 122 A; 126 C; 359 G; 400 T; 0 other;
 SQ
 Query Match 4.7%; Score 84.4; DB 24; Length 1007;
 Best Local Similarity 68.2%; Pred. No. 3.5e-12;
 Matches 146; Conservative 0; Mismatches 66; Indels 2; Gaps 2;

DT 12-JUL-2002 (first entry)
 XX
 XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 31746.
 DE
 XX
 XX Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 XX Homo sapiens.
 OS
 XX WO200218632-A2.
 PN
 XX
 PD 07-MAR-2002.
 XX
 XX 01-SEP-2001; 2001WO-EP10074.
 PF
 XX 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K, Guetig D:
 PI
 DR WPI; 2002-371829/40.
 XX
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 CC
 XX
 XX Sequence 1007 BP; 400 A; 359 C; 126 G; 122 T; 0 other;
 SQ
 Query Match 4.7%; Score 84.4; DB 24; Length 1007;
 Best Local Similarity 68.2%; Pred. No. 3.5e-12;
 Matches 146; Conservative 0; Mismatches 66; Indels 2; Gaps 2;

DB 889 GAAAGAA 883

RESULT 12
ABN81171

ID ABN81171 standard; DNA; 170 BP.

XX ABN81171;

DE 16-JUL-2002 (first entry)

XX Shrimp polynucleotide TUDGLV3-5.289F.

XX Giant black tiger prawn; Penaeus monodon; pacific white shrimp;

XX Litopenaeus vannamei; shrimp; microsatellite sequence; genome mapping;

XX Taura Syndrome Virus; TSV; infection; ds.

XX Litopenaeus vannamei.

XX WO200034476-A2.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US29571.

XX 10-DEC-1998; 98US-0111670.

XX (TUFTS) TUFTS COLLEGE.

XX Alciwar-Warren A, Xu Z, Dhar AK, Fan Y, Meehan D, Garcia DK;

XX WPI; 2000-423422/36.

XX Polynucleotides of shrimp are useful for identifying, mapping and

XX characterizing of the genome of various species of shrimp -

XX Claim 1, Fig 1; 120pp; English.

XX The invention relates to an isolated polynucleotide (1) of the giant

XX black tiger prawn, Penaeus monodon or expressed sequence tags of the

XX pacific white shrimp, Litopenaeus vannamei (ABN80997-ABN81172), both

XX containing microsatellites sequences including those P. monodon

XX complementary sequence or fragment and the encoded polypeptide are useful

XX for mapping of the genome of various species of shrimp. Mapping the

XX genome of Penaeus is useful for determining whether a test shrimp,

XX preferably Litopenaeus vannamei, has a genotype associated with a

XX phenotypic trait such as resistance to Taura Syndrome Virus (TSV)

XX infection.

XX Sequence 170 BP; 78 A; 11 C; 73 G; 8 T; 0 other;

XX SQ

Query Match 3.2%; Score 58.4; DB 21; Length 170;

Best Local Similarity 63.6%; Pred No. 1.3e-05;

Matches 89; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

DB 334 TGTATTAAGGAGCTGTATGATATGAGGGAAGCTGGACAGGGAAGGAGAG 393

DB 31 TATATTAAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 90

DB 394 GACGAG 453

DB 91 AG 150

AC ABN81079;

XX 16-JUL-2002 (first entry)

XX Shrimp polynucleotide SEQ ID NO 83.

XX Giant black tiger prawn; Penaeus monodon; pacific white shrimp;

XX Litopenaeus vannamei; shrimp; microsatellite sequence; genome mapping;

XX Taura Syndrome Virus; TSV; infection; ds.

XX Penaeus monodon.

XX WO200034476-A2.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US29571.

XX 10-DEC-1998; 98US-0111670.

XX (TUFTS) TUFTS COLLEGE.

XX Alciwar-Warren A, Xu Z, Dhar AK, Fan Y, Meehan D, Garcia DK;

XX WPI; 2000-423422/36.

XX Polynucleotides of shrimp are useful for identifying, mapping and

XX characterizing of the genome of various species of shrimp -

XX Claim 1, Page 97; 120pp; English.

XX The invention relates to an isolated polynucleotide (1) of the giant

XX black tiger prawn, Penaeus monodon or expressed sequence tags of the

XX pacific white shrimp, Litopenaeus vannamei (ABN80997-ABN81172), both

XX containing microsatellites sequences including those P. monodon

XX complementary sequence or fragment and the encoded polypeptide are useful

XX for mapping of the genome of various species of shrimp. Mapping the

XX genome of Penaeus is useful for determining whether a test shrimp,

XX preferably Litopenaeus vannamei, has a genotype associated with a

XX phenotypic trait such as resistance to Taura Syndrome Virus (TSV)

XX infection.

XX Sequence 502 BP; 81 A; 162 C; 41 G; 218 T; 0 other;

XX SQ

Query Match 3.2%; Score 57.8; DB 21; Length 502;

Best Local Similarity 59.4%; Pred No. 3.2e-05;

Matches 98; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

DB 320 AAGGAGATATATATATATTAAGGAGAGCTGTATGATATGAGGGAAGAGAGAGG 379

DB 428 AATGATATATATGCTGTATGATATATATGATATATGAGAGAGAGAGAGAGAGAGG 369

DB 380 GGAAG 439

DB 368 AG 309

DB 440 CGAGAGATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484

DB 308 AG 264

RESULT 14

AAAF64796

ID AAFA64796 standard; cDNA; 396 BP.

XX AAFA64796;

XX 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 552.

XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;

KW breast cancer; lung cancer; cancer detection; ss.
XX Homo sapiens.
XX WO200102568-A2.
XX 11-JAN-2001.
XX 30-JUN-2000; 2000WO-US18374.
XX 02-JUL-1999; 99US-0142310.
XX 02-JUL-1999; 99US-0142311.
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
DR WPI; 2001-091805/10.
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX Claim 9; Page 623-624; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 396 BP; 111 A; 81 C; 121 G; 79 T; 4 other;
Query Match 3.2%; Score 57.4; DB 22; Length 396;
Best Local Similarity 57.1%; Pred. No. 3.6e-05;
Matches 100; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 289 AGAGGAGCAAAATGATCTTCCATTCTAAAGAGAGATATATATATAAAGAGAC 348
DB 10 AG 69
QY 349 TGTATATGATATGGGGAAGAGTGACAGGGGGAAGAGAGAGAGAGAGAGAG 408
DB 70 GAG 139
QY 409 GGGAG 463
DB 130 GAG 184

RESULT 15
AAF67104
ID AAF67104 standard; cDNA; 420 BP.

AA67104;

09-APR-2001 (first entry)

Novel human polynucleotide, SEQ ID NO: 2860.

XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX Homo sapiens.
XX WO200102568-A2.
XX 11-JAN-2001.
XX 30-JUN-2000; 2000WO-US18374.
XX 02-JUL-1999; 99US-0142310.
XX 02-JUL-1999; 99US-0142311.
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
DR WPI; 2001-091805/10.
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX Claim 9; Page 970; 1046pp; English.
XX
XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 420 BP; 115 A; 80 C; 132 G; 93 T; 0 other;
Query Match 3.2%; Score 57; DB 22; Length 420;
Best Local Similarity 63.5%; Pred. No. 4.8e-05;
Matches 87; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 340 AAGGAAAGCGTATATATGGGGAAGAGTGACAGGGGGAAGAGAGAGAGAG 399
DB 12 AG 71
QY 400 AG 459
DB 72 AG 131
QY 460 ACAGAAATGATGATGA 476
DB 132 AG 148

Search completed: February 23, 2003, 08:33:24
Job time : 466 secs

COMMENT

Web : www.genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using Chinese BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutomo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source	location/qualifiers			
	1..932	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"			
	/clone="BACR14B09"			
	/clone_1ib="RPCI-98"			
	/note="end : T7"			

BASE COUNT	155 a	202 c	241 g	91 t	243 others
ORIGIN					

Query Match	4.0%; Score 71.6; DB 17;	Length 932;
Best Local Similarity	34.8%; Pred.No.7,2e-06;	
Matches 114;	Conservative 73;	Mismatches 141; Indels 0; Gaps 0;
OY	1477 GATCGGGGCTGTGAGTTGCCAGAGTACTGGGCTGCACCGGGTGCTGAGTGGCCT	1536
Dd	494 GKGKGGGGGGGVCGRGGSGSGSTGGCGGGCCCKCTSGGTGGTGGTGGCGG	553
OY	1537 GGAGTGCAGCGGGGTGGCCGACACTTGTCTCTGTAGACACAAGAAGTAGAGGGCTCG	1596
Dd	554 GSKSGSGSGSGSBCCGGGGSTTKGSGSGGSBSPTKTYBKRGKGGCGSSGSCS	613
OY	1597 GCGGTGAGACAGAGACAGCAAGTGTCCCAGGCGAGTCTCAGCTGCACAGCCCCGAGCTGG	1656
Dd	614 GGGGGCGSGGGGGCGGGGSSCSGCSGCGSGGGCGCCGSSGSGSGSGSGSG	673
OY	1657 CCCCGCTGTCAGGGGCGAAGCACGGGGCCCCGACCAGCCGCTTGGGCGGACCCGCCCCC	1716
Dd	674 SGGCGGSGCGCGSSGCCSCCCCGCGCGSGCGCGGSCGCGSGSGSGSGSG	733
OY	1717 TCACAACCCCACTCTCTGGGCGCGGCTTCCGGGGGCGGTCTCTGGGCCAACCCCGGCTTCA	1776
Dd	734 SCSCCCSCCCSSSCGSSSGCGSGGSGCGSCGCGSGCGSGCGSGGSGSGSS	793
OY	1777 TATAGGGCGCGCGCGCGCGCGCGCGCC	1804
Dd	794 GCGSGSSGSGCGSSSGSSCGGGSGCSS	821

RESULT 4	BMB10045	1364 bp	mRNA	linear	EST 05-MAR-2007
LOCUS	BMB10045				
DEFINITION	AGNCDCORT.6579545 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:5469986				
ACCESSION	BMB10045				
VERSION	BMB10045.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eumetazoa; Metazoa; Chordata; Craniota; Vertebrate; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 1364)				
AUTHORS	NIH-MGC http://mhc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-tr@mail.nih.gov				

	BASE COUNT	40 a	356 c	673 g	27 t	268 others
ORIGIN						
Query Match						3.9%; Score 70.8; DB 14; Length 1364;
Best Local Similarity						47.7%; Pred. No. 1.1e-05;
Matches 173; Conservative						0; Mismatches 189; Indels 1; Gaps 1;
OY 1442	GAGACGGGAAGGCGGTAAAGCGCGCCGGGGTCAAGAATCGGGGTCTGATTGCCAAG	1501				
DB 605	GG	664				
OY 1502	ACTGACTGGGGTGACCGGGGGGTCGTAGAGTGCCCTGAGAGTCCCGGGTTGCCGGCACAC	1561				
DB 665	GGGGNNGGGGCGCCGNNNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	724				
OY 1562	CTTGGTCTTGTAGACGACAAGTAGACGGGCTCCGGCGGTGACGACGAGACAGGTGC	1621				
DB 725	GG	784				
OY 1622	CGGGCAGTCTCGAGCTCAGCGCCCCCGAGTCCGGCCCGCGTGTCTCAAGGCGAAGCACG	1681				
DB 785	CGGG	844				
OY 1682	GGCCCCCGCAGCCGTGCTCGCCCGCACCCGGCCCGCTCCCAACC-CGACTCTCTGGGGCGG	1740				
DB 845	CGCGNGGG	904				
OY 1741	CGTTCGGGGGGGTGCTCTGGGGCACCCCGCGCTTATAATACGGGGCGGGCGCGGGCC	1800				
DB 905	CGCCCCGG	964				
OY 1801	GCC 1803					
DB 965	GNC 967					
RESULT 5						
LOCUS	BM914930					
DEFINITION	AGNENCOURT 6607405 NIH_MGC_41 Homo sapiens cdna clone IMAGE:5480653					
VERSION	BM914930					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					

RESULT	7
BQ434910/c	
LOCUS	
DEFINITION	BQ434910 1021 bp mRNA linear EST_24-MAY-2002 AGNNCCOURT_7909708 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6156444 5' , mRNA sequence.
ACCESSION	BQ434910
VERSION	BQ434910.1
KEYWORDS	GI:21173986
SOURCE	EST.
ORGANISM	human. Homo sapiens Euarthropoda; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1021) NIH-MGC http://mgc.ncl.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D.
JOURNAL	
COMMENT	

FEATURES	Location/Qualifiers
source	1. .1021

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAMA13500 row: 1 column: 13
 High quality sequence stop: 333.

BASE COUNT	55 a	353 c	463 g	48 t	102 others
ORIGIN					

Query Match	3.9%	Score 69.8;	DB 14;	Length 1021;
Best Local Similarity	46.8%;	Pred. NO. 1.9e-05;		
Matches 170; Conservative	0;	Mismatches 193;	Indels 0;	Gaps 0;

QY	1442	GAGACGCGGAAAGGGGTAAAGCGCGCGGGGGGTCAAGAGATATCGGGGGTGTCTAGTTGGCCAGG	1501
Db	773	GGGGGGCCCGCGGGCGGGGGGNCCTCCGCGGGGGGCGCGGGGCCCGGGGCGGGGCGGG	714
QY	1502	AGTACGTGGGGGTACCGGGGGTGTCTAGGTGGCTTGGATGGCGGGGGTGGCCGGGACAC	1561
Db	713	NNGGGGGGGGCGGNCGGGGGGCGCGGGCGGCGCCCGCGCGCGGGGGCGGGCGGGGGCGC	654
QY	1562	CTTGGTTCTTTGTAAGACGACAAAGGTGACGGGGCTCCGGCGCTGAGCAGAGAGACAGTGC	1621
Db	653	CCCCGGGGGGCGCGCGGGGGGCGCCGAGGCGGGGGCGCGGCGCCCGCGGGGGGGCGG	594
QY	1622	CGGGGAGGTCTGACGTCTGACAGCGCCCCCGGAGCTCGGGCTGTGCTCAGGGCGAAGACAG	1681
Db	593	CGGGGGCCCCGGGGGGGCGCCCGGGCGGAGCGGGCGCCGCTCCGCTCCGCCCCCGGGCGCC	534
QY	1682	GAGCCCGCAGACCTGTGCTGTGCGAGACCGCGCCCGCTCCCAACCGCCACCTCTGGGGCGAG	1741
Db	533	CCCCCCCCCGCCGNCNCCCGCCCCCGCCCCCGCCCCCGCCCCCGCCGNCNCCGNCNCCCGCC	474
QY	1742	GTTTCGGGGGGGTGTCTGGGGCCACCCCGGCTTCTATATACGGGGCGGGCGCGCGCGGCGG	1801
Db	473	CCCCCCCCCGCCCCCGCCGNCNCCCGCCGNCNCCGNCNCCGNCNCCGNCNCCCGCCCCCGCC	414
QY	1802	CCC 1804	
Db	413	CCC 411	

RESULT	8				
LOCUS	BM562099				
DEFINITION	BM562099	1273 bp	mRNA	linear	EST 20-PEB-2002
ACCESSION	AGNCNCOURT_6597561				
VERSION	5', mRNA sequence.				
KEYWORDS	BM562099				
SOURCE	BM562099.1	GI:18807897			
ORGANISM	human.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 1273)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Wellcome Sanger Institute				

FEATURES

cdna library preparation: Rubin Laboratory
cdna library arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
plate: L1CM2004 row: b column: 13
high quality sequence step: 206.

BASE COUNT	96 a	450 c	537 g	33 t	157 others
ORIGIN					

Query Match	3.8%;	Score 69.2;	DB 13;	Length 1273;
Best Local Similarity	35.6%;	Pred. No. 2.5e-05;		
Matches 179; Conservative	0;	Mismatches 324;	Indels 0;	Gaps 0

QY	1297	GAGACCCAAATGTAAAGCAGAAATGTGCCAAGACTTCGGCTGCCAAGCAATTAATTTT	1356
Db	193	GAGACTGCNCCNN	252
QY	1357	AGAAAGCTCCACGAGGTACACAGGAATGCGGAGCGCTGATGCCAGTTTCCCGACACCG	1418
Db	253	NN	312
QY	1417	GCTCGCCCGCAGGAGACCTCACCCCGAGACGGTAAGCGCGCGCGGGCTCAAGA	1476
Db	313	NN	372
QY	1477	GATCGGGGTGCTGAGTTGGCCAGAGATGACTGGGGGTGCTGAGTGGCT	1536
Db	373	GCGGGGGGGGGGGCCNCGGGGGGGGGGGGNNCGCGGGGGGGGGGGGGGGGGGGGGGG	432
QY	1537	GGAATGCCCGGGGTGGCGCGGCACACTTGTGTTTAAACACAAGTGAAGGGCTCCG	1596
Db	433	GCGGG	492
QY	1597	GCGCTGAGCAGAGAGACAGTTCGCCGGCGAGTTCGAGCTGCACGCCGCCCGAGCTCGG	1656

Dd	493	G G G G G G G G G G G G G G G G C C G C C G C C G G G G G G G G G G C C C C C C G G G G G G C C C G G C C C G G G C G C	552
Oy	1657	C C C C G G C T C T C A G G G C G A G C A G C G G G C C C C C A G C G C T C T G G C C G C A C C C G G C C C C	1716
Dd	553	C G G G C C G C G G C C C G G G C C G G C C G G G G C C C C C C G C C C C C G G C C C C C C C C C C G G G C G G C	612
Oy	1717	T C C C A C C C C C A C T C C T G G G C G C G G C T T C C G G G G C G T G T C T G G G C A C C C C G G C T T C T A	1776
Dd	613	G G C C G G C C C C C C C C C C C C C C C C C C G C G G C G G C A C C C G G G G C C C G C G G C C C	672
Oy	1777	T A T A C G G G G C G G C G G C C C G G G C	1799
Dd	673	G C C G G G G G G C C G G G C C G G G G G G C	695

RESULT	9
CNS006XK	
LOCUS	CNS006XK
DEFINITION	CNS006XK 935 bp DNA linear GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION	BACR1409 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL066051
VERSION	AL066051.1 GI:4945019
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 935)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT

Determination of this BAC-end sequence was carried out as part of collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammasser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPT1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	Location/qualifiers
SOURCE	1. .935

BASE COUNT ORIGIN	257 a	170 c	162 g	96 t	250 others
----------------------	-------	-------	-------	------	------------

Query Match	3.8%	Score 68.8	DB 17	Length 935
Best Local Similarity	29.1%	Pred. No. 3.2e-05		
Matches 124	Conservative 116	Mismatches 184	Indels 2	Gaps 1

1381 AATGGGAGCGCTGTATGCCAGTTTCCCCGACACCGGCTCGCCGAGGAGACTCACCC 1440

Db 503 DMMVCGGSCSMCRTSSSGTGCSTCCMYMSSSVSCCSCSGTCCGYCSSCTSCMKMCCTYG 562

Db 563 CKCGCGCSTSCSSCCSBBSYSTCCTBCTKCSSGCTGSGCTGCCGGGGGCGCG 622

[illegible]

RESULT	10
B0673485	
LOCUS	1190 bp mRNA
DEFINITION	AGSCNCOURT 8189102 NIH_MGC_102 Homo sapiens cDNA IMAGE:6254865
ACCESSION	B0673485
VERSION	B0673485
KEYWORDS	EST.
SOURCE	human.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1190)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Plate: L1CM2405 row: a column: 10
High quality sequence stop: 226.
Location/Qualifiers

source	1. .1190
--------	----------

```

//tissue_type="epidermoid carcinoma, cell line"
//lab_host="DH10B (phage-resistant)"
//note="Organ: salivary gland; Vector: pOTB7; Site_1: xhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/xhoI sites using the
following 5' adaptor: GGCACGAC(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NH2MGC Library."

```

Query Match	3.7%	Score 67.6;	DB 14;	Length 1150;
Best Local Similarity	50.6%	Pred. No. 5.8e-05;		

3

2

RESULT 11	CNS004NB	839 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	CNS004NB				
DEFINITION	Drosophila melanogaster genome survey sequence Tm13 end of BAC #				
	BAC1016 of Rpel-38 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				

ACCESSION	AL054280
VERSION	AL054280.1
KEYWORDS	GI:4931788
SOURCE	GSS.
ORGANISM	Drosophila melanogaster.
	Drosophila melanogaster.
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 839)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT

Bp 191 91006 EVRY cedex - FRANCE (E-mail: secref@genomecope.cns.fr)
- Web : www.genomecope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library was prepared by Kazuoyo Oseegawa and Aaron Mammosoer in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

1. .839

```
/organism="Drosophila melanogaster"  
/db_xref="taxon:7227"
```

```
/clone="BACR10E16"  
/clone_lib="RPCI-98"
```

BASE COUNT	/note="end : TEJ3"			
ORIGIN	285 a	67 c	77 g	27 t
				383 others

Query Match 3.7%; Score 67.2; DB 17; Length 839;

QY	1392	CTGTATCTCCAGT	TTCCCGGCACACCGGCTGC	CGCCAGAGAGACTTCACCCGGAAGCGAA	1451
		↑	↑	↑	↑
Db	441	CNNATATATNTNT	TAATAAGAAATATSMSSSSACASAS	SSSSCGASSCASSCVAACASACAGSSAS	500
QY	1452	GGGGTAAAGCGCGCGGGGCTCAAGGAGATCGGGGTCGTGATTTGGCCACAGAGTCACTGG	1511		
		↑	↑	↑	↑
Db	501	AGSSSSSAGSCGAGSSSSSAGSGSAGASGSGSAGASGSGSGSGRSGSCGAGVASAGS	560		
QY	1512	GTGACGGGGGGTCTGAGGNGGCTGTGAGGCGCTGTGAGCGCGGGCGCGGGCACACCTTGGTTCTT	1571		
		↑	↑	↑	↑
Db	561	SSACSCASSASSSSSASGCMSCSSSSSSSAAACSSSSSSSVAAVCGSGSGVSPA	620		
QY	1572	GTACACGACAGGTGACGGGGCTCCGGGCGGTGAGCAAGAGACAGATGCCGGGCGAGTCTC	1631		
		↑	↑	↑	↑
Db	621	AAVARGARGCMGMAAGSGRAYSAASAASASASASVAGACSSGSGAGSSSAGAA	680		
QY	1632	TCGAGCTGACACGCCCCCGGACTGTGGGCCCGCGCTGCTCAGGCGGAAACACAGGCCCCCGCAG	1691		
		↑	↑	↑	↑
Db	681	RARASVCAASASBSGSGSGSSSMCAVASBSGAASBSAGCGCGSSSQCSCSSSM	740		
QY	1692	CCGTGCTGTGGCGGACCGCGGCCCTTCCCAACCCCCACACTCTGTGGGGCGCGTGTTCGGGGGC	1751		
		↑	↑	↑	↑
Db	741	SCSSSSCGSSSGCVCSCSCSVSCSCSCBBSBSCSCCASSSRASGCGSSSSSCSSVCGSM	800		

QY 1752 GTGTCCTGGGCCACCCCG 1769
: | ::|:::
Db 801 SGCMCAMSSASSASSSS 818

RESULT 12	CNS0091P	LOCUS	DEFINITION
	CNS0091P	925 bp	DNA linear GSS 03-JUN-1999
			Drosophila melanogaster genome survey sequence TE13 end of BAC #
			BACR19D16 of RPECI-98 library from Drosophila melanogaster (fruit
			fly), genomic survey sequence.

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AL053013	AL053013.1	GI:4934461	GSS.	<i>Drosophila melanogaster</i> .	
				<i>Drosophila melanogaster</i>	
				Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota	
				Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
				Phytridae; Drosophilidae; <i>Drosophila</i> .	
				1 (bases 1 to 925)	

AUTHORS
TITLE
JOURNAL

COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genome.cns.fr
- Web : www.genoscope.cns.fr)

determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oseegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .925

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_1lb="RPCI-98"

```


BASE COUNT	71 a	373 c	538 g	102 t	78 others
ORIGIN					

Query Match	3.7%;	Score 66.8;	DB 14;	Length 1162;
Best Local Similarity	49.2%;	Pred. No. 8.9e-05;		
Matches 178; Conservative	0;	Mismatches 183;	Indels 1;	Gaps 1

QY	1442	GAGACCGGAAAGGGGTAAAGGCAGCGCGGGCTCAAGAGATCAGGAGTTCAGTTGGCCAGG	1501
Db	473	GGCGGGGGGGGNGGGGGGGGGGGGGGGGCGNCGGGGGGGGGGGGGGGGGGGGGGGGGGG	532
QY	1502	AGTGACTGGGGGTGACCGGGGGCTGTGAAGGTGCTTGAGTGC CGGGGGTGGCCGGCACAC	1561
Db	533	GGGGGGGGGNGGGGGGGGGGGGGGGGGGGGGGNGGGGGGGGGGGGGGGGGGGGGGGGG	592
QY	1562	CTTGTTCTTGTAGACGACAAGTACCGGCTCCGGGGCTGTAGCAACGAGACACAGTGGC	1621
Db	593	GGGGGGGGGGCGGGGGGGGNGNNGGGNGGGGGGGGGGGGGGGGGGGGGGCCGGGGGGGG	652
QY	1622	C GG G G A G T C T C A G A C T C A C A G C C C C C G A G C T C G G A C C C G G T G C T C A G A G C G A A G A C A G	1681
Db	653	GGCGCGGGCCCCCGGGCGCCCGCCCCCGGGCGGCGAGGGGGCCCCCGCGCGGGGGCCGG	712
QY	1682	GAGCCCGACGCGCTGCTGCGCGCGAACCCGCCCTCTCCACACCOCACACTCTCTGGCGCGC	1741
Db	713	GAGCGCGGCGCGGGCGCCCGCCGC -CGGGCCCCCCNCNMCGCCCGGGCGCCGGGGGGGGG	771
QY	1742	GTTCCGGGGGCGGTGTCCTTGCGGCCACCCCGGGCTTCTATTATCGGGCGCGCGCGCCGGCCG	1801
Db	772	GCGCCCGCCCGCGCGCGCGCCCGCCCGCGCGCGCGCCCGCCCGCGCGCGCGCGCGCGCG	831
QY	1802	CC 1803	
Db	832	CC 833	

LOCUS	B0943816				
DEFINITION	B0943816	1136 bp	mRNA	linear	EST 21-AUG-2007
ACCESSION	AGENCOURT_8773466	NIH_MGC_18	Homo sapiens	CDNA clone	IMAGE:6372516
VERSION	B0943816				
KEYWORDS	B0943816.1	GI:22359294			
ORGANISM	EST.				
SOURCE	human.				
REFERENCE	Homo sapiens				
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
JOURNAL	1 (bases 1 to 1136)				
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .				
	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapds@email.nih.gov				
	Tissue Procurement: DCTD/DTF/Gazdar				
	CDNA Library Preparation: Rubin Laboratory				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium (LLNL)				
	at: http://image.llnl.gov				
	Plate: L1CM2548 row: g column: 13				
	High quality sequence start: 54				
	High quality sequence stop: 323.				

FEATURES

source

ECORI; cDNA made by oligo-dT priming. Directionally cloned
 into ECORI/XhoI sites using the following 5' adaptor:
 GGCGAGGAG(c). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC library.

Query Match	3.7%	Score	66.6	DB	14	Length	1136
Best Local Similarity	47.1%	Pred. NO.	9.9e-05				
Matches 171; Conservative	0	Mismatches	192			Indels	0
						Gaps	0

QY	1442	GAGACCGGAAGGGCTTAAAGCGCGCCGGGGTCAAGAGATCGGGGGTCTAGATTGGCCAGG	1501
Db	544	GGGGGGGGGGGGGGGGGGNNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	603
QY	1502	AGTGACTTGGGGGTGAACCGGGGGTGTGAGGTGGCTTGAAGTGGCCGGGGTGGCCGGCACAC	1561
Db	604	GGGGGNGGGGGGGGGGGGNGGGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	663
QY	1562	CTTGGTTCTTGTAGACGACAAAGGTGACGGGCTTCCGGGCGTGAACACGAGGAGCAAGTGCC	1621
Db	664	GGGGGGGGGGGGGGGGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	723
QY	1622	CGGGGAGTGTGTCGAGCTGACAGCGCCCGCGAGCTGGGCGCCGGGGTGTCAAGGGGGAAGCAAG	1681
Db	724	GGGGGGGGGGGGGGGGGGGGGGGGCCCGACMGGGGGGGGGGGGGGGGGGGGGGGGGGG	783
QY	1682	GAGCCCGCCAGACCGTGGCTGTGGCGGACCCGCGCCCTCCCAACCCCACTCTGGGCGGCG	1741
Db	784	GGCGCCCGGG	843
QY	1742	GTTTCCGGGGCGTGTCTTGGGCCAACCCCGCGTTCATATATACGGGCGGGCGCGCCGGGGCGG	1801
Db	844	NCGGGGCGGGCGCGCCCGGNCGCCCGGAGGCGCCCGCGCGACGCGCGCGCGGGGGGG	903
QY	1802	CCC	1804
Db	904	CGC	906

QY	1442	GAGACCGGAAGGGCTTAAGCGCGCGCGGGTCAAGAGATCGGGGGTCTAGATTGGCCAGG	1501
Db	544	GGGGGGGGGGGGGGGGGGNNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	603
QY	1502	AGTGACTTGGGGGTGAACCGGGGGTGTGAGGTGGCTTGAAGTGGCGGGGGTGGCCGACAC	1561
Db	604	GGCGGNGGGGGGGGGGGGNGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	663
QY	1562	CTTGCGTTCTTGTAGACGACAAAGGTGACGGGCTTCCGGGCGTGAACACGAGGAGCAAGTGC	1621
Db	664	GGGGGGGGGGGGGGGGGGNMGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	723
QY	1622	CGGGCGAGTGTGCGAGTGCAGCGCCGCCCGAGCTGGGCGCCGGGCGTCAAGGGGGAAGCAAG	1681
Db	724	GGGGGGGGGGGGGGGGGGGGGGGGGCCCGACMGGGGGGGGGGGGGGGGGGGGGGG	783
QY	1682	GAGCCCGCCAGACCGTGCCTGTGGCGGACCCGCGCCCTCCCAACCCCACTCTGGGCGGC	1741
Db	784	GGCGCCCGGG	843
QY	1742	GTTTCGGGGCGTGTCTTGGGCCAACCCCGCGTCTATATACGGGCGGGCGCGCCGGGGCGG	1801
Db	844	NCGGGGCGGGCGCGCCCGGNCGCCCGGAGGCGCCCGCGCGACGCGCGCGGGGGGG	903
QY	1802	CCC	1804
Db	904	CGC	906

Search completed: February 23, 2003, 10:31:25
Job time : 2400 secs

[illegible]

```

RESULT 2
US-08-253-155A-23
; Sequence 23, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeno
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 329 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-253-155A-23

```

	Query Match	3.1%	Score 56.2	DB 1	length 329
	Best Local Similarity	61.1%	Pred. No. 8.7e-06		
	Matches	91	Conservative	0	Mismatches 95
					Indels 0
					Gaps 0
Qy	340	AAAGCAACTGTATATGATATGTGGGGCAAGAGGTGGACAGGGGGAAAAAGGCGACAGCACGAG	399		
Db	25	AGAGGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG	84		

OY	400	AGAGAGAAAGGAGGAGGAGGAGCACAAGAGAGCACTGGGCGAGAGATCGATTACGAGAG	455
Db	85	AGAGAGGAGAGAGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	140
OY	460	ACAGAAATGATGATGAGAAATTAACCTTA	488
Db	145	AGAGAGAGCTCTATGATCTTTCCATTTCA	173

RESULT 3
 US-07-891-942G-11
 : Sequence 11, Application US/07891942G
 : Patent No. 5679511
 : GENERAL INFORMATION:
 : APPLICANT: Kwon, Byoung Se
 : TITLE OF INVENTION: CDNA CLONES FOR HUMAN TYROSINASE AND FOR
 : TITLE OF INVENTION: A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY
 : NUMBER OF SEQUENCES: 16
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Christopher A. Michaels, Barnard, Brown &
 : ADDRESSEE: Michaels
 : STREET: 306 East State Street, Suite 220
 : CITY: Ithaca
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 14850
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/891,942G
 : FILING DATE: 01-JUN-1992
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 06/915,753
 : FILING DATE: 06-OCT-1986
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/362,847
 : FILING DATE: 07-JUN-1989
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Michaels, Christopher A
 : REGISTRATION NUMBER: 34,390
 : REFERENCE/DOCKET NUMBER: INDI
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 607-273-1711
 : TELEFAX: 607-273-2609
 : INFORMATION FOR SEQ ID NO: 11:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2397 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : OS-07-891-942G-11

[illegible]

1 APPLICANT: Gyuris, Jeno
 2 APPLICANT: Draetta, Giulio
 3 TITLE OF INVENTION: CDK4 Binding Proteins
 4 NUMBER OF SEQUENCES: 95
 5 CORRESPONDENCE ADDRESS:
 6 ADDRESSEE: LAHIVE & COCKFIELD
 7 STREET: 60 State Street
 8 CITY: Boston
 9 STATE: MA
 10 COUNTRY: USA
 11 ZIP: 02109
 12 COMPUTER READABLE FORM:
 13 MEDIUM TYPE: floppy disk
 14 COMPUTER: IBM PC compatible
 15 OPERATING SYSTEM: PC-DOS/MS-DOS
 16 SOFTWARE: ASCII(text)
 17 CURRENT APPLICATION DATA:
 18 APPLICATION NUMBER: US/08/253,155A
 19 FILING DATE: 02-JUN-1994
 20 CLASSIFICATION: 435
 21 ATTORNEY/AGENT INFORMATION:
 22 NAME: Vincent, Matthew P.
 23 REGISTRATION NUMBER: 36,709
 24 REFERENCE/DOCKET NUMBER: MIT-028
 25 TELECOMMUNICATION INFORMATION:
 26 TELEPHONE: (617) 227-7400
 27 TELEFAX: (617) 227-5941
 28 INFORMATION FOR SEQ ID NO: 14:
 29 SEQUENCE CHARACTERISTICS:
 30 LENGTH: 349 base pairs
 31 TYPE: nucleic acid
 32 STRANDEDNESS: single
 33 TOPOLOGY: linear
 34 MOLECULE TYPE: cDNA
 35 US-08-253-155A-14
 36 Query Match 2.9% Score 53; DB 1; Length 349;
 37 Best Local Similarity 58.1% Pred. NO. 6.5e-05;
 38 Matches 93; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 340 AAGAGAGCTGTATGATATGAGGAGAGAGTGCACAGGGGAGAAAGGAGAGAGCAG 399
DB 51 AG 110
QY 400 AG 459
DB 111 AG 170
QY 460 ACAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 499
DB 171 AG 210

RESULT 7
US-08-306-691B-14
Sequence 14, Application US/08306691B
Patent No. 5734039
GENERAL INFORMATION:
APPLICANT: Calabretta, Bruno
APPLICANT: Skorski, Tomasz
TITLE OF INVENTION: ANTISENSE
TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P.C.
STREET: Two Penn Center, Suite 1800
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5734039e
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 6453 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-306-691B-14

Query Match 2.9%; Score 51.6; DB 1; Length 6453;
Best Local Similarity 46.2%; Pred. No. 0.00063;
Matches 171; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

DB 721 TGAGCGGCGCTGCGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 790
QY 1586 GACGGGCTCCGGGCGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645
DB 791 TCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850
QY 1646 CCCGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1705
DB 851 CTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 910
QY 1706 ACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1765
DB 911 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 970
QY 1766 CCGGCGGCTCT 1775
DB 971 AGTGGCTTTT 980

RESULT 8
US-09-209-668-10
Sequence 10, Application US/09209668A
Patent No. 6114517
GENERAL INFORMATION:
APPLICANT: Monia, Brett P.
TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
FILE REFERENCE: ISPH-0336
CURRENT APPLICATION NUMBER: US/09/209,668A
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 6453
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1664)..(1774)
FEATURE:
NAME/KEY: CDS
LOCATION: (2042)..(2220)
FEATURE:
NAME/KEY: CDS
LOCATION: (2374)..(2533)
FEATURE:
NAME/KEY: CDS
LOCATION: (3231)..(3350)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: V00574/Genbank
DATABASE ENTRY DATE: 1991-01-03
US-09-209-668-10

Query Match 2.98%; Score 51.6; DB 3; Length 6453;
Best Local Similarity 46.2%; Pred. No. 0.00063;
Matches 171; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

```

QY 1646 CCGAGCTCGGCGCCGCTGCTGAGGGCGAAGACGCGGCCCGACGCTGCTGCGCG 1705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 CTGCGGCGCCGCGCGCCCTACACACCGGGGGGCTGAGGAGGAGCGCGCGCGACGCGA 910
QY 1706 ACCCGCGCCCGCTCCGACCCCGCACTGCTGGGCGGCGCTTCCGGGCGCTGCGCGAC 1765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CGCGCGGCGACCCCGCATGACATCACAGGTGCGGAGCACGCGCGGCGCTCAGCCCC 970
QY 1766 CCGGCGCTCT 1775
    ||| |||
Db 971 AGTGCTTTT 980

```

```

RESULT 9
US-09-952-8
; Sequence 8, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borjick-Stodin, Ann
; APPLICANT: Margalit, S. M.
; APPLICANT: Bor-Sogil, Dafna
; APPLICANT: Cole, Phillip
; APPLICANT: Kurtyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; EARLIER FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/099,631
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-356-952-8

```

```

Query Match
Best Local Similarity 46.2%; Score 51.6; DB 3; Length 6453;
Matches 171; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

```

```

QY 1406 CCGGAGACCGGCTCGCGCGGAGAGACTTCACCCGAGAGCGGAAGGGTAAAGGCGGC 1465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 611 CTCTCTCGCGCCCGAGGCGAGCAATACGCGCGCGCGCGCGCGCGCGCGCGCGCG 670
QY 1466 GGGGTCAAGAGATCGGGGCTGCTGAGTTGCCAGAGTACTGGGCTGACCGGGGCTGC 1525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 GGGGTAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 730
QY 1526 TGAGGTGAGCTGAGTCCGCGGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 TGAGCGGGGCTGCGGGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAA 790
QY 1586 GAGCGGCTCGCGGCTGAGACGAGAGAGAGTGCCCGCGGAGTCTGAGCTGCGACGCC 1645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 TCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 850
QY 1646 CCGGAGCTCGGCGCGCGGCTGCTCAGGGCGAAGACGCGCGCGCGCGCGCGCGCGCG 1705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 851 CTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 910
QY 1706 ACCCGCGCCCGCTCCCAACCCGCACTCTTGGGCGCGCGCTTCCGGGCGCTGCTGAGCGAC 1765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CGCGCGGCGACCCCGCATGATCAGATCAGAGTGGCGGAGCACGCGCGCGCGCGCTCAGCCCC 970
QY 1766 CCGGCGCTCT 1775
    ||| |||
Db 971 AGTGCTTTT 980

```

RESULT 10

```

US-09-570-367C-1/c
; Sequence 1, Application US/09570367C
; Patent No. 6338851
; GENERAL INFORMATION:
; APPLICANT: Gorczynski, Reginald M.
; TITLE OF INVENTION: Methods and Compositions for Immunomodulation
; FILE REFERENCE: 9579-21
; CURRENT APPLICATION NUMBER: US/09/570,367C
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/064,764
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentln Version 3.0
; SEQ ID NO 1
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-570-367C-1

```

```

Query Match
Best Local Similarity 57.5%; Score 51.2; DB 4; Length 2791;
Matches 92; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 318 AAAAGCAGTATATATATATATAAGAGAGCTGTATGATATGCGGCGAGAGTGACAG 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 AGACAGAAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAAAG 323
QY 378 GGGGAAAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 263
QY 438 GCGGAGATCGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 AAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223

```

```

RESULT 11
US-08-772-440-33
; Sequence 33, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Aritizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DCTIN-1 AND DCTIN-2; COMPOSITIONS AND USES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10409 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified_base
LOCATION: 6510
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "D - A or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3406..6470
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "K - G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3564..7896
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "M - A or C"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3497..3607
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "N - A or C or G or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3479..6422
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "R - A or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3405..6871
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "S - C or G"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3457..9998
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "W - A or T"
FEATURE:
NAME/KEY: modified_base
LOCATION: 3595..9999
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /note- "Y - C or T"
US-08-772-440-33

Query Match 2.8%; Score 50.4; DB 3; Length 10409;
Best Local Similarity 50.4%; Pred. No. 0.0017;
Matches 123; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 226 CATGTTCTGCTTGGGCTCTTTTATTTAGAGGTCCTCATTTATTTATTTAGT 285
DB 4904 CTGTTGTTCTCTGTCAGTTCAGTCTGCTGTTCTTCTTTTAAATTTT 4963
QY 286 ACAAGGAGGAGCAAAATGTTCTTTCAATCTAAAGAGAGTATATATATAAAGGA 345
DB 4964 AAATTTGTCTTTAAAAAATTTGCCCCGTTATTTTCAAGAGAGAGAGAGAGAG 5023
QY 346 AGCTGTATGATATGGGAGAGAGTGGACAGGGGAAAAAGGAGAGAGAGAGAG 405
DB 5024 AG 5083
QY 406 AAAGGAGGAGAGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
DB 5084 AG 5143
QY 466 ATGA 469
DB 5144 TGA 5147

RESULT 12
US-08-810-572A-1
; Sequence 1, Application US/08810572A

Patent No. 5969102
GENERAL INFORMATION:
APPLICANT: Bram, Richard J.
APPLICANT: von Bulow, Gotz
TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,572A
FILING DATE: 28-FEB-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1377 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-810-572A-1

Query Match 2.8%; Score 50.2; DB 2; Length 1377;
Best Local Similarity 62.2%; Pred. No. 0.00071;
Matches 79; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 346 AGCTGTATGATATGGGAGAGAGTGGACAGGGGAAAAAGGAGAGAGAGAGAG 405
DB 1082 AGAGGAAAGAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
QY 406 AAAGGAGGAGAGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
DB 1142 AGAGGAG 1201
QY 466 ATGATGA 472
DB 1202 AGCAGGA 1208

RESULT 13
US-09-290-333-1
; Sequence 1, Application US/09290333
Patent No. 6316222
GENERAL INFORMATION:
APPLICANT: Bram, Richard J.
APPLICANT: von Bulow, Gotz
TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/290,333
FILING DATE: 12-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1340-1-007 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1377 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-290-333-1

Query Match 2.8%; Score 50.2; DB 4; Length 1377;
Best Local Similarity 62.2%; Pred. No. 0.00071;
Matches 79; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 346 AGCTGTATAGATATGCGGCGAGAGTGGACAGGCGGAAAGGAGAGAGAGAGAG 405
DB 1082 AGAGGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1141
QY 406 AAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 465
DB 1142 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1201
QY 466 ATGATGA 472
DB 1202 AGCAGGA 1208

RESULT 14
US-08-658-136-2/c
Sequence 2, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA

ZIP: 01701

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 53526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-2

Query Match 2.8%; Score 50; DB 3; Length 53526;
Best Local Similarity 49.2%; Pred. No. 0.0047;
Matches 131; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1538 GAGTGGCGGGGTGGCGGGGACACCTTGTGTTAGACAGCAAGGTGACGGGCTCGG 1597
DB 3887 GGGTGGCGGGGTGGCGGGGACACCTTGTGTTAGACAGCAAGGTGACGGGCTCGG 3828
QY 1598 GCGTGACAGAGAGAGAGAGTGGCGGGGCGAGTCTCGAGCTGCACGCCCGGACTGGC 1657
DB 3827 ACCGAGGCTCGGACGCCCGGGGCGAGCACTTACCGGAGCGGCGGCGGCGCTGG 3768
QY 1658 CCGGCTGCTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1717
DB 3767 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3708
QY 1718 CCGAAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTAT 1777
DB 3707 CCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3648
QY 1778 ATAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1803
DB 3647 CGTTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3622

RESULT 15
US-08-658-136-1/c
Sequence 1, Application US/08658136
Patent No. 6071717
GENERAL INFORMATION:
APPLICANT: KLINGER, KATHERINE W
APPLICANT: LANDES, GREGORY M
APPLICANT: BURN, TIMOTHY C
APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/658,136
;
; FILING DATE:
;
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: LASSEN, ELIZABETH
;
; REGISTRATION NUMBER: 31,845
;
; REFERENCE/DOCKET NUMBER: GEN4-17.8
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 508-872-8400
;
; TELEFAX: 508-872-5415
;
; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 53577 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
; US-08-658-136-1

```

```

Query Match      2.88; Score 50; DB 3; Length 53577;
Best Local Similarity 49.2%; Pred. No. 0.0047;
Matches 131; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1538 GAGTGGCGGGGTGGCGGCGACACCTGTGTTGTAGACGACAAGTGACGGGCTCGG 1597
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3887 GGGTGGCGGCTGGGCGGCGCTACTACAGCGCTGTGGCTCCGCGGATGCGCAGCGGG 3828
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1598 GCGTGAGCACAGAGAGAGAGTGCCCGGCGAGTCTGAGCTGACAGCGCCCGAGCTCGG 1657
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3827 ACCGAGCGTCCGAGCGCGCGGCGCGAGCAATTGACGGGCGAGCGGCGCGGCGCTGG 3768
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1658 CCGGCGTCTCAGAGGCGACAGCGGCGCCCGCAGCGCTGCTGCGCCGACCGCCCT 1717
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3767 GCGCGAGAGCGAGGCGGCTCGCAGGGCCGCGCGCGCGGCGCGCCCGCGCAGCGC 3708
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1718 CCCAACCCCTCTCTGCGGCGGCTTCGCGGCGGTGCTGCTGAGGCAACCCCGCTTCTAT 1777
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3707 CCGAGCCACAGGCGCCAGGCGCCAGGCGCCAGCGCGGCGGCGCGCGCGCGCAT 3648
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1778 ATACGGGCGCGCGCGCGCGCGCC 1803
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3647 CGTAGGCGAGCGCGCGCATGGCCCC 3622
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: February 23, 2003, 10:37:08
 Job time : 353 secs

Qy	1	GGGAGTTCGTTTTGTTTTCCTTCATTTTCCCAATAAATACAGAGATCCGTCTTGAGGTG	60
Db	101	GGGGATTCGTTTTGTTTTCCTTCATTTTCCCAATAAATACAGAGATCCGTCTTGAGGTG	166
Qy	61	CAAGCAGATCTCTGAAGAGAGATGATACAGAGAAAG-ATACAGCAAGCAGCATTTATGG	119
Db	161	CAAGCAGATCTCTGAAGAGAGATGATACAGAGAAAGAAATACAGCAAGCATTTATGG	220
Qy	120	CACGGTTTCTGTAAACAGGTTGAGTGTAGCCACAGCTGTAGCACTGTGTGGAGAAAGAGC	179
Db	221	CACGGTTTCTGTAAACAGGTTGAGTGTAGCCACAGCTGTAGCACTGTGTGGAGAAAGAGC	280


```
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (390986)..(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (346860)..(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (317174)..(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (280353)..(280373)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (271829)..(271848)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (183872)..(183891)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (170625)..(170645)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (132680)..(132700)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
```

```
Query Match 3.6%; Score 65.6; DB 9; Length 659158;
Best Local Similarity 63.1%; Pred. No. 3.7e-06;
Matches 101; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
```

```
QY 318 AAAGAGAGTATATATGATAAAGAGCTGTATAGATATGAGGAGAGGTGGACAG 377
||||| ||| ||||| | ||| ||| ||| ||| |||
DB 463003 AAAAGAGAGATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 462944
QY 378 GGGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 437
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 462943 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462884
QY 438 GCGAGAGATCGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 462883 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462844
```

RESULT 3
US-09-771-208-20

```
; Sequence 20, Application US/09771208
; Patent No. US2002015564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFOR, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
```

```
NAME/KEY: misc_feature
LOCATION: (546998)..(547017)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (494715)..(494814)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (390986)..(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (346860)..(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (317174)..(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (280353)..(280373)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (271829)..(271848)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (183872)..(183891)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (170625)..(170645)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
LOCATION: (132680)..(132700)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc_feature
OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
```

```
Query Match 3.2%; Score 58.6; DB 9; Length 659158;
Best Local Similarity 64.2%; Pred. No. 0.0003;
Matches 88; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
QY 340 AAAGAGAGCTGTATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343768 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343827
QY 400 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343828 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343887
QY 460 ACAGAGAGATGATGATGA 476
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 343888 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343904
```

RESULT 4

```
US-09-938-842A-3578/c
; Sequence 3578, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: HARPER, JEFF
; APPLICANT: KREPS, JOEL
; APPLICANT: WANG, XUN
; APPLICANT: ZHU, TONG
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3578
```



```

?       TOPOLOGY: linear
?       MOLECULE TYPE: genomic DNA
?       HYPOTHETICAL: NO
?       ANTI-SENSE: NO
?       FRAGMENT TYPE: <unknown>
?       ORIGINAL SOURCE:
?       SEQUENCE DESCRIPTION: SEQ ID NO: 17
US-03-799-462A-17

```

Query Match	2.9%	Score 52.8	DB 9	Length 42999
Best Local Similarity	61.8%	Pred. No. 0.0022		
Matches 84	Conservative 0	Mismatches 52	Indels 0	Gaps 0

QY	342	AGGAAGCTGTATTAGTATTATGGGGAGAGAGTGTGACACAGGGGGAAAAGGGGAGACAGCAGAG	401
Db	14947	AGGGGAGGAGGAGGAGGCGAGACAGCAAGACAGAGAGAGACAGAGAGCAGCAGAGAGAG	14888
QY	402	AGACAAAGGAGGAGGAGGAGGACCAAGAGAGAGCACTATGGCGAGAGATTCGATTAGGAGAGAC	461
Db	14887	AGAGAAAGAGAGAGAGAGAGGCGGAGCCGAGCAGAGACAGAGAGAGAGAGAGAGAGAGAGAG	14828
QY	462	AGAAATCATCATCA	477
Db	14827	AGAGAAAGAGAAAGAA	14812

RESULT 8
 US-10-125-767-17/c
 : Sequence 17, Application US/10125767
 : Patent No. US20020160410A1
 : GENERAL INFORMATION:
 : APPLICANT: Hadlaczky, Gyula
 : TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND
 : METHODS
 : FOR PREPARING ARTIFICIAL CHROMOSOMES
 :
 : NUMBER OF SEQUENCES: 34
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Heller Erhman White & McAnuliffe LLP
 : STREET: 4350 La Jolla Village Drive, 7th Floor
 : CITY: San Diego
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92122
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSEO Version 1.5
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/10/125,767
 : FILING DATE: 17-Apr-2002
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 09/724,693
 : FILING DATE: 28-NOV-2000
 : APPLICATION NUMBER: 08/835,682
 : FILING DATE: 10-APR-1997
 : APPLICATION NUMBER: 08/695,191
 : FILING DATE: 07-AUG-1996
 : APPLICATION NUMBER: 08/682,080
 : FILING DATE: 15-JUL-1996
 : APPLICATION NUMBER: 08/629,822
 : FILING DATE: 10-APR-1996
 :
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Seidman, Stephanie L.
 : REGISTRATION NUMBER: 33,779
 : REFERENCE/DOCKET NUMBER: 24601-402J
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 858-450-8403
 : TELEFAX: 858-587-5360
 : TELEX: <Unknown>
 :
 : INFORMATION FOR SEQ ID NO.: 17:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 42999 base pairs

```
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: Genomic DNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE: <unknown>
? ORIGINAL SOURCE:
? SEQUENCE DESCRIPTION: SEQ ID NO: 17
?-IS-10-125-/67-17
```

Query Match	2.9%	Score 52.8	DB 9	Length 42999
Best Local Similarity	61.8%	Pred. No. 0.0022		
Matches 84	Conservative 0	Mismatches 52	Indels 0	Gaps 0

OY	342	AGAAAGCTGTATAGATTATGGCGGGAAGAGTGTGACAGGGGGGAAAAAGGCGAGACGACGAGAG	401
Dd	14947	AGGGGAGGAGGAGGAGGAGAGAGAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	14888
OY	402	AGAGAAAAGGAGAGGAGAGGAGCAAGAGAGACTGTGGCGAGAGAGATTCGATTAGGAGAGAC	461
Dd	14887	AGAGAAAAGGAGAGAGAGAGGCGCGAGACCGAGACAGAGAGAGAGAGAGAGAGAGAGAG	14828
OY	462	AGAAATGATGATGAATGAA	477
Dd	14827	AGAGAAAAGGAAAAAGAA	14812

RESULT 9
 US-09-836-911A-17/c
 Sequence 17, Application US/09836911A
 Publication No. US2003003617A1
 GENERAL INFORMATION:
 APPLICANT: Hadlaczky, Gyula
 Szalay, Aladar
 TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
 AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Heller Ehrman White & McCulliffe
 STREET: 4350 La Jolla Village Drive, 6th Floor
 CITY: San Diego
 STATE: CA
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/836,911A
 FILING DATE: 17-Apr-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/835,682
 FILING DATE: 10-APR-1997
 APPLICATION NUMBER: 08/695,191
 FILING DATE: 07-AUG-1996
 APPLICATION NUMBER: 08/682,080
 FILING DATE: 15-JUL-1996
 APPLICATION NUMBER: 08/629,822
 FILING DATE: 10-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 24601-4021
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 858-450-8403
 TELEFAX: 858-587-5360
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:

PRIOR FILING DATE: 2001-01-30

QY 460 ACAGAAATGATGAA 473

STIC-Biotech/ChemLib

86812

Fr m: Yu, Misook
Sent: Friday, February 14, 2003 9:22 AM
To: STIC-Biotech/ChemLib
Subject: 09/865,879

RECEIVED

FEB 14 2003

STIC-BIOTECH/ChemLib
(STIC)

Please search SEQ ID NO:2

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

TYPE OF SEARCH:

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/20
Date Completed: 2/24
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:17:17 ; Search time 4104 Seconds

(without alignments)
11012.831 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaacc.....gttcgcgtcgtccgcctcc 1553Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1553	100.0	1553	6	AX323410	AX323410 Sequence
2	1542	99.3	76007	9	AC0004503	AC0004503 Homo sapi
3	105	6.8	2748	9	AK094581	AK094581 Homo sapi
4	96.6	6.2	7218	6	I66494	I66494 Sequence 14
5	75.8	4.9	207420	2	AC078884	AC078884 Mus muscu
6	73	4.7	151283	9	AC107540	AC107540 Homo sapi
7	71	4.6	2583	2	AK094055	AK094055 Homo sapi
8	70.2	4.5	86057	2	AC131553	AC131553 Homo sapi
9	65.4	4.2	100740	2	AC131216	AC131216 Rattus no
10	65	4.2	2659	9	AK093916	AK093916 Homo sapi
11	65	4.2	119989	2	AC119472	AC119472 Rattus no
12	62.4	4.0	132222	2	AC121192	AC121192 Rattus no
13	62.4	4.0	147622	2	AC110437	AC110437 Rattus no
14	61.8	4.0	201650	2	AC119304	AC119304 Rattus no
15	61.8	4.0	203343	2	AC131133	AC131133 Rattus no
16	61.6	4.0	194562	2	AC112394	AC112394 Rattus no
17	61.4	4.0	69051	2	AC073171	AC073171 Homo sapi
18	61.2	3.9	52683	2	AC131481	AC131481 Rattus no
19	61	3.9	131956	2	AC131361	AC131361 Rattus no
20	61	3.9	190626	2	AC098181	AC098181 Rattus no
21	61	3.9	206773	2	AC126296	AC126296 Rattus no
22	60.8	3.9	106649	2	AC122626	AC122626 Rattus no
23	60.8	3.9	196828	2	AC127709	AC127709 Rattus no
24	60.6	3.9	182269	2	AC107416	AC107416 Rattus no
25	60.6	3.9	232603	2	AC127560	AC127560 Mus muscu
26	60.2	3.9	69636	2	AC129731	AC129731 Rattus no
27	60.2	3.9	112289	2	AC125732	AC125732 Rattus no
28	60	3.9	141288	2	AC116264	AC116264 Rattus no
29	59.8	3.9	204431	2	AC096927	AC096927 Rattus no
30	59.2	3.8	173022	2	AC128405	AC128405 Rattus no
31	59.2	3.8	17801	2	AC114391	AC114391 Rattus no
32	59	3.8	2790	6	AX281844	AX281844 Sequence
33	58.6	3.8	121077	2	AC105558	AC105558 Rattus no
34	58.6	3.8	152347	2	AC021096	AC021096 Homo sapi
35	58.4	3.8	143786	2	AC112363	AC112363 Rattus no
36	58.4	3.8	197949	2	AC109737	AC109737 Rattus no
37	58.2	3.7	132884	2	AC128403	AC128403 Rattus no
38	58	3.7	150305	2	AC114705	AC114705 Rattus no
39	58	3.7	16641	2	AC126956	AC126956 Rattus no
40	58	3.7	182225	2	AC128440	AC128440 Rattus no
41	58	3.7	186906	2	AC110640	AC110640 Rattus no
42	57.8	3.7	105620	2	AC127760	AC127760 Rattus no
43	57.8	3.7	133052	2	AC112614	AC112614 Rattus no
44	57.8	3.7	140340	2	AC108584	AC108584 Rattus no
45	57.6	3.7	145592	2	AC119341	AC119341 Rattus no

ALIGNMENTS

RESULT 1
AX323410
LOCUS AX323410 1553 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 2 from Patent WO0192578.
ACCESSION AX323410
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Roninson, I.B., Dokmanovic, M. and Chang, B.D.
TITLE Reagents and methods for identifying and modulating expression of
genes regulated by retinoids

JOURNAL Patent: WO 0192578-A 2 06-DEC-2001;
Board of Trustees of the University of Illinois (US)
FEATURES Location/Qualifiers
source 1. .1553
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Beta IG-H3 promoter NCBI acc. number AC004503.1"

BASE COUNT 305 a 490 c 384 g 374 t
ORIGIN

Query Match 100.0%; Score 1553; DB 6; Length 1553;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACTCAGGAGTGTCCCAACCACTATCTACCTGGCAAGCTGCACTGTGCTGCTCA 60
1 GACTCAGGAGTGTCCCAACCACTATCTACCTGGCAAGCTGCACTGTGCTGCTCA 60
61 TTCTGAACATGGAGCACTACTGTGCAATGTCCAGACACAAACACCTTCAATATCT 120
61 TTCTGAACATGGAGCACTACTGTGCAATGTCCAGACACAAACACCTTCAATATCT 120
121 TGACTCTCTCTCTCCCTCTCTCTCTATACAGACTCAAAATTTGAGACTATTAC 180
121 TGACTCTCTCTCTCCCTCTCTCTCTATACAGACTCAAAATTTGAGACTATTAC 180
181 CTCTTACACCCCTTCACTATTTGCCAGCTTCCCATCTGTGCTCAACCATATGTTCA 240
181 CTCTTACACCCCTTCACTATTTGCCAGCTTCCCATCTGTGCTCAACCATATGTTCA 240
241 AGCTCCCAATGGTCCCT 300
241 AGCTCCCAATGGTCCCT 300
301 CACTGGCAATGTACTTGTCTCCACCACTACACCGCTAGGCTTTAGCTGAGTCTGGC 360
301 CACTGGCAATGTACTTGTCTCCACCACTACACCGCTAGGCTTTAGCTGAGTCTGGC 360
361 CACTGGCAATGTACTTGTCTCCACCACTACACCGCTAGGCTTTAGCTGAGTCTGGC 360
361 CACTGGCAATGTACTTGTCTCCACCACTACACCGCTAGGCTTTAGCTGAGTCTGGC 360
421 CGAGATTAATCTGATTTGCGAATGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
421 CGAGATTAATCTGATTTGCGAATGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
481 TTTCAGGCTCTAGAGAGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 540
481 TTTCAGGCTCTAGAGAGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 540
541 AGCCCAAGCTACTCTTCTATGGAATCATGATGCTGATGCTGATGCTGATGCTGATGCT 600
541 AGCCCAAGCTACTCTTCTATGGAATCATGATGCTGATGCTGATGCTGATGCTGATGCT 600
601 GCTTTTTCACAAAGGCTTGAAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 660
601 GCTTTTTCACAAAGGCTTGAAGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 660
661 GCTAGAGTCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
661 GCTAGAGTCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 TACCATGTGACAAATTTGTTTACCTGCTTGAAGTCTGATGCTGATGCTGATGCTGAT 780
721 TACCATGTGACAAATTTGTTTACCTGCTTGAAGTCTGATGCTGATGCTGATGCTGAT 780
781 CAGATCTATCTGCGAGTACTGTTGCGAGATGTTTACTGAGAGAGAGATGATGATGAT 840
781 CAGATCTATCTGCGAGTACTGTTGCGAGATGTTTACTGAGAGAGAGATGATGATGAT 840
841 ACCAAGGTGAG 900
841 ACCAAGGTGAG 900

901 ACCTGCCCCACATTTTGGCGTTTCTGTGATCAGACAGATTTACTCATCTGTTCAAT 960
901 ACCTGCCCCACATTTTGGCGTTTCTGTGATCAGACAGATTTACTCATCTGTTCAAT 960
961 GAGCATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1020
961 GAGCATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1020
1021 CAGCTCTGGGGTGCAGAAACAAAGTGTGATGGGCAATGTTCTGGGAATCAGACTGC 1080
1021 CAGCTCTGGGGTGCAGAAACAAAGTGTGATGGGCAATGTTCTGGGAATCAGACTGC 1080
1081 CTCCCTACCCCTCTCTGAGTGTGTTTGTAGCTCAAGATCAGTGAAGGAATCTTGGG 1140
1081 CTCCCTACCCCTCTCTGAGTGTGTTTGTAGCTCAAGATCAGTGAAGGAATCTTGGG 1140
1141 CCCCCAGATGACAG 1200
1141 CCCCCAGATGACAG 1200
1201 TTGGAACAGACAG 1260
1201 TTGGAACAGACAG 1260
1261 GAGGAG 1320
1261 GAGGAG 1320
1321 CACAGTGGGAG 1380
1321 CACAGTGGGAG 1380
1381 AGGCGGCGCGCCCTCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1381 AGGCGGCGCGCCCTCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1441 TCCGCTGCGAGCTTACTTAACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
1441 TCCGCTGCGAGCTTACTTAACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
1501 CCGCGCGCTTGGCG 1553
1501 CCGCGCGCTTGGCG 1553

RESULT 2
AC004503
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC004503 76007 bp DNA linear PRI 30-MAR-1998
Homo sapiens chromosome 5, p1 clone 1354A7 (LBNM, H47), complete
sequence.
AC004503 AC002223 AC002225 AC001513 AC002226 AC001514
AC002222 AC001516 AC001515 AC001511 AC001510 AC001512 AC002221
AC002224
AC004503.1 GI:2996635
HTG.
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 76007)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
Sequencing of human chromosome 5
Unpublished
2 (bases 1 to 76007)
Ricke,D.O.
Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
Unpublished
3 (bases 1 to 76007)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,

TITLE Davis, C.A., Kadner, K., Miguel, T., Pitluck, S., Pollard, M.,
JOURNAL Rojsek, H., Sudramanian, S. and Martin, C.H.
Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES Location/Qualifiers
1. 76007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q"
/clone="1354A7"
/note="LBNL H47"
complement(1. 39820)
/note="99% identity AC002428"
repeat_region complement(992. 1315)
/rpt_family="Alu"
repeat_region 1355. 1444
/rpt_family="MSTC"
repeat_region 1558. 1705
/rpt_family="MSTB"
complement(2377. 5032)
/rpt_family="L1"
repeat_region complement(6287. 6744)
/rpt_family="Alu"
complement(8102. 9558)
/rpt_family="L1"
misc_feature complement(9587. 9700)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 12179. 12494
/rpt_family="MLT1a"
repeat_region 15197. 15345
/rpt_family="MLT1a"
repeat_region 15425. 15597
/rpt_family="MLT1a"
complement(17298. 17716)
/rpt_family="L1"
repeat_region complement(18631. 18813)
/rpt_family="MLT1e"
repeat_region complement(18989. 19025)
/rpt_family="MLT1"
repeat_region 19607. 19636
/note="(GT)15"
/rpt_type=tandem
/rpt_unit=GT
misc_feature complement(23378. 23472)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(23777. 24117)
/rpt_family="Alu"
repeat_region 24513. 24586
/rpt_family="MER68B"
repeat_region complement(2785. 27918)
/rpt_family="MIR2"
misc_feature complement(28314. 28444)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 29788. 30276
/rpt_family="Alu"
repeat_region 30242. 30272
/note="(A)31"
/rpt_type=tandem
/rpt_unit=A
repeat_region 30482. 30800
/rpt_family="MER33"
repeat_region complement(32562. 33906)
/rpt_family="L1"
repeat_region complement(35700. 35884)
/rpt_family="MER5"
repeat_region complement(36451. 36650)
/rpt_family="MIR"
repeat_region complement(37382. 37458)

repeat_region /rpt_family="MIR2"
complement(41227. 41588)
/rpt_family="THE1"
repeat_region 41633. 41763
/rpt_family="MIR"
misc_feature 42050. 42204
/note="GRAIL 2 excellent exon, frame 0"
44455. 44591
/note="GRAIL 2 excellent exon, frame 2"
complement(46368. 46472)
/rpt_family="MIR"
repeat_region 48908. >75715
/gene="BIGH3"
/note="transforming growth factor-beta induced gene
product (BIGH3)"
join(48955. 49088, 53662. 53760, 63957. 64021, 66234. 66394,
66750. 66914, 67173. 67319, 69338. 69479, 72806. 73018,
73842. 73979, 74615. 74760, 75579. >75715)
/gene="BIGH3"
/note="transforming growth factor-beta induced gene
product (BIGH3)"
/codon_start=1
/evidence=experimental
/product="BIGH3"
/protein_id="AAC08449.1"
/db_xref="GI:2996636"
/translation="MALFVRLALALALGPAATLAPAKSPYOLVLRGRH
GNVAVOKVIGNKRKYFTNGKQWQKICGKSTVSYSCGGEKVPKCPALP
LSNLTETGVSSTGTYDTRERKPRMERGSTTTPASNEANASIPAYLDSLV
SVNTEILNALRHVGRVYLDLHGMFTLSMKNNSNIQIHHPNGLVTVNCALC
KADHATNGVAVHIDKVISFTNNIQQIIEIDVETTLAAVAASGLNMLRNGCYT
LIAPNEAPEKIPSETLNRILGDPBALRLNNHILKSAACAAIYAGLSVELEGT
LEVGCSGDMLTNGKAIISNKDILATNGVYIHIDELIPDAKTELELAESQVST
DLEPQGLGNHLSGSEKRLIAPINSGVFGDPPIDAHNRLLRNHITDOLASKYI
HGQLETTGSKRLRVYRNSLCIENSCIAHDKRRTGLTFMDRVLPKGTADV
LKGDNR"

repeat_region 55224. 55552
/rpt_family="Alu"
repeat_region 58141. 58334
/rpt_family="MLT1"
repeat_region 58303. 58632
/rpt_family="Alu"
repeat_region 58629. 58886
/rpt_family="MLT1"
repeat_region 60208. 60533
/rpt_family="MER1B"
complement(62232. 62307)
/rpt_family="MIR"
63480. 63559
/rpt_family="MIR"
complement(63619. 63756)
/note="GRAIL 2 excellent exon, frame 0"
complement(64678. 64950)
/rpt_family="MER2"
65024. 65474
/rpt_family="Alu"
65447. 65474
/note="(AAAT)7"
/rpt_type=tandem
/rpt_unit=AAAT
repeat_region 69938. 70262
/rpt_family="Alu"
71056. 71368
/rpt_family="Alu"
71445. 71590
/gene="BIGH3"
/note="GRAIL 2 excellent exon, frame 0"
71652. 71691
/gene="BIGH3"
/note="GRAIL 2 excellent exon, frame 0"
75793. 76002
/note="100% identity B43847 BAC end sequence"
BASE COUNT 20153 a 16885 c 17588 g 21381 t

ORIGIN

Query Match 99.3%; Score 1542; DB 9; Length 76007;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1553; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GACTCAGGCTGCCAACAACACATCATCTACCTGGCAAGCTGCACCTGCATGTCCTCA 60
 |||||||
 DB 47401 GACTCAGGCTGCCAACAACACATCATCTACCTGGCAAGCTGCACCTGCATGTCCTCA 47460
 QY 61 TTCTGAACATGGACACATCA-CTGCTGCAATGTCAGACACAAACACCCACAAATATCC 119
 |||||||
 DB 47461 TTCTGAACATGGACACATCACTGCTGCAATGTCAGACACAAACACCCACAAATATCC 47520
 QY 120 TTGACT 179
 |||||||
 DB 47521 TTGACT 47580
 QY 180 CCTCTACACCCCTCACATTTGGCCAGCCCTTCCCAATCTGCTCTACACCATATGTC 239
 |||||||
 DB 47581 CCTCTACACCCCTCACATTTGGCCAGCCCTTCCCAATCTGCTCTACACCATATGTC 47640
 QY 240 AAGCTCTCCATGCTGCT 299
 |||||||
 DB 47641 AAGCTCTCCATGCTGCT 47700
 QY 300 ACACCTGACATGTCATCTGCT 359
 |||||||
 DB 47701 ACACCTGACATGTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 47760
 QY 360 CCTCTGACCT 419
 |||||||
 DB 47761 CCTCTGACCT 47820
 QY 420 ACCAGATTAATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
 |||||||
 DB 47821 ACCAGATTAATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 47880
 QY 480 CTTTCAGGCTCAGACAGATATCACCTCTGTCAGAGGCTGCTCTCTCTCTCTCTCTCTCA 539
 |||||||
 DB 47881 CTTTCAGGCTCAGACAGATATCACCTCTGTCAGAGGCTGCTCTCTCTCTCTCTCTCTCA 47940
 QY 540 AAGCCCGACCT 599
 |||||||
 DB 47941 AAGCCCGACCT 48000
 QY 600 AACTTTTACAAAGGCTTAGGCTCTACATCTCTATTAATCCATCTGTCCTCCAG 659
 |||||||
 DB 48001 AACTTTTACAAAGGCTTAGGCTCTACATCTCTATTAATCCATCTGTCCTCCAG 48060
 QY 660 GGCTACAGTGCAGATTAAGTGAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTA 719
 |||||||
 DB 48061 GGCTACAGTGCAGATTAAGTGAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTA 48120
 QY 720 TTACCATGTCACATTTGTTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCT 779
 |||||||
 DB 48121 TTACCATGTCACATTTGTTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCT 48180
 QY 780 GGAGATCTATCTGGAGGATGTTGACAGATGTTTACGAAAGGAGGAAATTAAG 839
 |||||||
 DB 48181 GGAGATCTATCTGGAGGATGTTGACAGATGTTTACGAAAGGAGGAAATTAAG 48240
 QY 840 TACCAAGTGAAGACCCACCTTCCCAAGGCTGCCAAGAGAGTGAAGCCCAACAA 899
 |||||||
 DB 48241 TACCAAGTGAAGACCCACCTTCCCAAGGCTGCCAAGAGAGTGAAGCCCAACAA 48300
 QY 900 AACCTGCCCCACATTTTGGCGTTTCTGTCGATTCACAGATTTACTCATCTGTCCTTCA 959
 |||||||
 DB 48301 AACCTGCCCCACATTTTGGCGTTTCTGTCGATTCACAGATTTACTCATCTGTCCTTCA 48360
 QY 960 TGAGCATGACAGTGGGCTGGGCTGGAGGATTAAGATTAAGGAGCTGGGAGGCTGG 1019
 |||||||
 DB 48361 TGAGCATGACAGTGGGCTGGGCTGGAGGATTAAGATTAAGGAGCTGGGAGGCTGG 48420

QY 1020 TCAGCTCTGGGGTGCAGAAACAAGTCTGATGGGCCATGCTGTTCTGGGAATCAGCACTG 1079
 |||||||
 DB 48421 TCAGCTCTGGGGTGCAGAAACAAGTCTGATGGGCCATGCTGTTCTGGGAATCAGCACTG 48480
 QY 1080 CCTCCCTCACCCTCTCCCTGAGTGTGTTGATGCTTCAAGATCAGTGAAGGAATCTTGG 1139
 |||||||
 DB 48481 CCTCCCTCACCCTCTCCCTGAGTGTGTTGATGCTTCAAGATCAGTGAAGGAATCTTGG 48540
 QY 1140 GCCCCAGCATGACGAGACCGAAGCCCGGAGACAGTGTCTCTCAGTCCCAAGTCCCA 1199
 |||||||
 DB 48541 GCCCCAGCATGACGAGACCGAAGCCCGGAGACAGTGTCTCTCAGTCCCAAGTCCCA 48600
 QY 1200 TTGGAAGCAGCCACAGAGGCTTAAGGACCTATACCTTGTGTTGAGGAAGACTGTGG 1259
 |||||||
 DB 48601 TTGGAAGCAGCCACAGAGGCTTAAGGACCTATACCTTGTGTTGAGGAAGACTGTGG 48660
 QY 1260 CGAGGAGACAGGAGGAGGCTGTCGTCAGAGGCAAGGCTGCGAAGAACTGACACGG 1319
 |||||||
 DB 48661 CGAGGAGACAGGAGGAGGCTGTCGTCAGAGGCAAGGCTGCGAAGAACTGACACGG 48720
 QY 1320 GCACAGTGGGAGACGGGTGGGTGCCAGGCGACAGGCGGCGAGGGTTGGAGAGCGC 1379
 |||||||
 DB 48721 GCACAGTGGGAGACGGGTGGGTGCCAGGCGACAGGCGGCGAGGGTTGGAGAGCGC 48780
 QY 1380 CAGGCGCGCGCCCTCTCTCTCAGCGGCGCGCCAGCTTCCCGCCCTGCGCTCC 1439
 |||||||
 DB 48781 CAGGCGCGCGCCCTCTCTCTCAGCGGCGCGCCAGCTTCCCGCCCTGCGCTCC 48840
 QY 1440 CTCCCGCTCGACGTTACTTAACCTGGCCGCGCGCGGAGGCGCTCTCACTTCCCTGA 1499
 |||||||
 DB 48841 CTCCCGCTCGACGTTACTTAACCTGGCCGCGCGGAGGCGCTCTCACTTCCCTGA 48900
 QY 1500 GCGCGCGCTGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1553
 |||||||
 DB 48901 GCGCGCGCTGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 48954

RESULT 3
 AK094581
 LOCUS
 DEFINITION
 Homo sapiens CDNA FLJ37262 f1s, clone BRAM2010919, highly similar
 to TRANSFORMING GROWTH FACTOR-BETA INDUCED PROTEIN IG-H3 PRECURSOR.
 ACCESSION
 AK094581
 VERSION
 AK094581.1 GI:21753668
 KEYWORDS
 oligo capping; f1s (full insert sequence).
 SOURCE
 Homo sapiens amygdala CDNA to mRNA, clone_1lb:BRAM2
 clone:BRAM2010919.

ORGANISM

Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
 Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
 Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
 Kawai, H., Otsu, K., Nishikawa, T., Kimura, K., Yamashita, H.,
 Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Matsunuma, M., Murakawa, K., Kanehori, K., Takahashi, Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahari, K., Masuno, T., Nagai, K. and Isogai, T.
 NEDO human CDNA sequencing project

Unpublished
 2 (bases 1 to 2748)

Isogai, T. and Yamamoto, J.
 Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human CDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan: cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

COMMENT


```
* 132177 137419: contig of 5243 bp in length
* 137420 137439: gap of unknown length
* 137440 142254: contig of 4815 bp in length
* 142255 142274: gap of unknown length
* 142375 149021: contig of 6747 bp in length
* 149022 149041: gap of unknown length
* 149042 152821: contig of 3780 bp in length
* 152822 152841: gap of unknown length
* 152842 157555: contig of 4714 bp in length
* 157556 157575: gap of unknown length
* 157576 160314: contig of 2739 bp in length
* 160315 160334: gap of unknown length
* 160335 164344: contig of 4210 bp in length
* 164345 164564: gap of unknown length
* 164565 166647: contig of 2083 bp in length
* 166648 166667: gap of unknown length
* 166668 169833: contig of 3166 bp in length
* 169834 169853: gap of unknown length
* 169854 173590: contig of 3737 bp in length
* 173591 173610: gap of unknown length
* 173611 178164: contig of 4554 bp in length
* 178165 180900: contig of 2716 bp in length
* 180901 180920: gap of unknown length
* 180921 182482: contig of 1562 bp in length
* 182483 182502: gap of unknown length
* 182503 185289: contig of 2787 bp in length
* 185290 185309: gap of unknown length
* 185310 186932: contig of 1623 bp in length
* 186933 186952: gap of unknown length
* 186953 188165: contig of 1213 bp in length
* 188166 188185: gap of unknown length
* 188186 189092: contig of 907 bp in length
* 189093 189112: gap of unknown length
* 189113 190678: contig of 1566 bp in length
* 190679 190698: gap of unknown length
* 190699 192877: contig of 2179 bp in length
* 192878 192897: gap of unknown length
* 192898 194214: contig of 1317 bp in length
* 194215 194234: gap of unknown length
* 194235 195535: contig of 1301 bp in length
* 195536 195555: gap of unknown length
* 195556 196407: contig of 852 bp in length
* 196408 196427: gap of unknown length
* 196428 197671: contig of 1244 bp in length
* 197672 197691: gap of unknown length
* 197692 198960: contig of 1269 bp in length
* 198961 198980: gap of unknown length
* 198981 200542: contig of 1562 bp in length
* 200543 200562: gap of unknown length
* 200563 201492: contig of 930 bp in length
* 201493 201512: gap of unknown length
* 201513 202368: contig of 856 bp in length
* 202369 202388: gap of unknown length
* 202389 203978: contig of 1590 bp in length
* 203979 203998: gap of unknown length
* 203999 205308: contig of 1310 bp in length
* 205309 205328: gap of unknown length
* 205329 207420: contig of 2092 bp in length.
```

FEATURES

SOURCE

```
1. 207420
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="17"
/clone="RP23-390F1"
/sex="male"
1. .63778
misc_feature
/note="assembly_name:Contig130"
63799. .96902
misc_feature
/note="assembly_name:Contig129"
96923. .107488
misc_feature
/note="assembly_name:Contig128"
```

```
misc_feature 107509. .119219
/note="assembly_name:Contig127"
misc_feature 119240. .132156
/note="assembly_name:Contig126
clone_end:T7
vector_side:right"
misc_feature 132177. .137419
/note="assembly_name:Contig125"
misc_feature 137440. .142254
/note="assembly_name:Contig124"
142275. .149021
/note="assembly_name:Contig123"
/note="assembly_name:Contig123"
149042. .152821
/note="assembly_name:Contig122"
152842. .157555
/note="assembly_name:Contig121"
157576. .160314
/note="assembly_name:Contig120"
160335. .164544
/note="assembly_name:Contig119"
164565. .166647
/note="assembly_name:Contig118"
166668. .169833
/note="assembly_name:Contig117
clone_end:Sp6
vector_side:right"
misc_feature 173611. .178164
/note="assembly_name:Contig116"
178185. .180900
/note="assembly_name:Contig115"
180921. .182482
/note="assembly_name:Contig114"
182503. .185289
/note="assembly_name:Contig113"
185310. .186932
/note="assembly_name:Contig112"
186953. .188165
/note="assembly_name:Contig111"
188186. .189092
/note="assembly_name:Contig110"
189113. .190678
/note="assembly_name:Contig109"
190699. .192877
/note="assembly_name:Contig108"
192898. .194214
/note="assembly_name:Contig107"
194235. .195535
/note="assembly_name:Contig106"
195556. .196407
/note="assembly_name:Contig105"
196428. .197671
/note="assembly_name:Contig104"
197692. .198960
/note="assembly_name:Contig103"
198981. .200542
/note="assembly_name:Contig102"
200563. .201492
/note="assembly_name:Contig101"
201513. .202368
/note="assembly_name:Contig99"
202389. .203978
/note="assembly_name:Contig98"
203999. .205308
/note="assembly_name:Contig97"
205329. .207420
```

```
Query Match 4.9%: Score 75.8; DB 2; Length 207420;
Best Local Similarity 52.9%; Pred No. 6 4e-08;
Matches 164; Conservative 0; Mismatches 144; Indels 2; Gaps 1;
Oy 1246 GAGGAGACGTGTGCGAGGAGAGAGGAGGAGGCGCTGCGAGTGAGGCGAGGCGCTGGC 1305
```


BASE COUNT	39694	a	32483	c	32177	g	40192	t	6737	others
ORIGIN										
Query Match	4.7%; Score 73; DB 2; Length 151283;									
Best Local Similarity	51.9%; Pred. No. 3.1e-07;									
Matches 154; Conservative	0; Mismatches 143; Indels 0; Gaps 0;									
OY 1248	GGAAGACTGTGGCGAAGGAGAAGAGGAGGAGGAGCTGCACGTGACGAGCCAGAGCGTGGGAA	1307								
Db 26762	GGGGGNNGGGGGGGGGGGNGGGGGGGGNGGGGGGGGGGGGGGNGGGGGGGGGGGGGGGGG	26821								
OY 1308	AACGACACAGGGCACAATGCCGGGAGCGGGTGTTGCCAGGGCAAGCCAGGGGGCGACGG	1367								
Db 26822	GG	26881								
OY 1368	GTTTGAGGAGGGGCGGAGCGGGCCCCCTTCCTTCACAGGGCGAGCCAGATTCCCAGCCCC	1427								
Db 26882	CAGCGGG	26941								
OY 1428	GGCGTCCGCTCCCTCCGCGTCCGACGATTAACCTGAGCCCGGGCGGAGGAGCGCTCT	1487								
Db 26942	GGCCCCGGCGGG	27001								
OY 1488	CACCTCCCTTGAGCGCCCGCTTGCCCGTGGCTGCTAGCTGCTCGGTGCGCGCTGG	1544								
Db 27002	CCGGCCCCCGGG	27058								
RESULT 7										
AKO94055										
LOCUS										
DEFINITION	Homo sapiens cDNA FLJ36736 fls, clone UTERU2013048, highly similar									
ACCESSION	AKO94055									
VERSION	AKO94055.1									
KEYWORDS	oligo capping; fts (full insert sequence).									
SOURCE	Homo sapiens uterus cDNA to mRNA, clone_lib:UTERU2									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
1	Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,									
	Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,									
	Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,									
	Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,									
	Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,									
	Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K.									
	and Isogai,T.									
TITLE	NEDO human cDNA sequencing project									
JOURNAL	Unpublished									
COMMENT	2 (bases 1 to 2583)									
REFERENCE	Isogai,T. and Yamamoto,J.									
AUTHORS	Direct Submission									
TITLE	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7									
JOURNAL	Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan									
COMMENT	(E-mail:igenomics@hri.co.jp, tel:81-438-52-3975, fax:81-438-52-3986									
	NEDO human cDNA sequencing project supported by Ministry of									
	Economy, Trade and Industry of Japan; cDNA full insert sequencing:									
	Research Association for Biotechnology (RAB); cDNA library									
	construction: Helix Research Institute (HRI) (supported by Japan									
	Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,									
	HRI, and Biotechnology Center, National Institute of Technology and									
	Evaluation; clone selection for full insert sequencing: HRI and									
	RAB; annotation: HRI and RAB.									
FEATURES	Location/Qualifiers									
source	1..2583									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/clone="UTERU2013048"									
	/tissue="uterus"									

TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

Newton, N., Nguyen, N., Norris, S., Nwackelam, O., Okunolu, G.,
Olanupinsagoh, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,
Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Tatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D.,
Walton, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R.,
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 100740)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (18-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: KAOA

Center clone name: CH230-62N13

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap, version 0.990329

Consensus quality: 47096 bases at least Q40

Consensus quality: 50680 bases at least Q30

Consensus quality: 54217 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 52 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1330: contig of 1330 bp in length
* 1331 1430: gap of unknown length
* 1431 2453: contig of 1023 bp in length
* 2454 2554: gap of unknown length
* 2554 3619: contig of 1066 bp in length
* 3620 3720: gap of unknown length
* 3720 5059: contig of 1340 bp in length
* 5060 5159: gap of unknown length
* 5160 6654: contig of 1495 bp in length
* 6655 6754: gap of unknown length
* 6755 7927: gap of 1172 bp in length
* 7927 8026: gap of unknown length
* 8027 9404: contig of 1378 bp in length
* 9405 9504: gap of unknown length
* 9505 10747: contig of 1243 bp in length
* 10748 10847: gap of unknown length
* 10848 12238: contig of 1391 bp in length
* 12239 12338: gap of unknown length
* 12339 13760: contig of 1422 bp in length
* 13761 13860: gap of unknown length
* 13861 15505: contig of 1645 bp in length

15506 15605: gap of unknown length
* 15606 17187: contig of 1582 bp in length
* 17188 17287: gap of unknown length
* 17288 18925: contig of 1638 bp in length
* 18926 19025: gap of unknown length
* 19026 20382: contig of 1357 bp in length
* 20383 20483: gap of unknown length
* 20483 21644: contig of 1162 bp in length
* 21645 21745: gap of unknown length
* 21745 23161: contig of 1417 bp in length
* 23162 23261: gap of unknown length
* 23262 24853: contig of 1552 bp in length
* 24854 24953: gap of unknown length
* 24954 25987: contig of 1034 bp in length
* 25988 26087: gap of unknown length
* 26088 27465: contig of 1378 bp in length
* 27466 27565: gap of unknown length
* 27566 29172: contig of 1607 bp in length
* 29173 29272: gap of unknown length
* 29273 30410: contig of 1138 bp in length
* 30411 30510: gap of unknown length
* 30511 32419: contig of 1909 bp in length
* 32420 32519: gap of unknown length
* 32520 34534: contig of 2015 bp in length
* 34535 34634: gap of unknown length
* 34635 36068: contig of 1434 bp in length
* 36069 36168: gap of unknown length
* 36169 37561: contig of 1393 bp in length
* 37562 37661: gap of unknown length
* 37662 39373: contig of 1712 bp in length
* 39374 39473: gap of unknown length
* 39474 41883: contig of 2410 bp in length
* 41884 41983: gap of unknown length
* 41984 43914: contig of 1931 bp in length
* 43915 44014: gap of unknown length
* 44015 45567: contig of 1553 bp in length
* 45568 45667: gap of unknown length
* 45668 47087: contig of 1420 bp in length
* 47088 47187: gap of unknown length
* 47188 48713: contig of 1526 bp in length
* 48714 48813: gap of unknown length
* 48814 50876: contig of 2063 bp in length
* 50877 50976: gap of unknown length
* 50977 52113: contig of 1137 bp in length
* 52114 52213: gap of unknown length
* 52215 53320: contig of 1107 bp in length
* 53321 53420: gap of unknown length
* 53421 55496: contig of 2076 bp in length
* 55497 55596: gap of unknown length
* 55597 56811: contig of 1215 bp in length
* 56812 56911: gap of unknown length
* 56912 58502: contig of 1591 bp in length
* 58503 58602: gap of unknown length
* 58603 59807: contig of 1205 bp in length
* 59808 59907: gap of unknown length
* 59909 62270: contig of 2363 bp in length
* 62271 62370: gap of unknown length
* 62371 64542: contig of 2172 bp in length
* 64543 64642: gap of unknown length
* 64643 66802: contig of 2160 bp in length
* 66803 66902: gap of unknown length
* 66903 70141: contig of 3239 bp in length
* 70142 70241: gap of unknown length
* 70242 72440: contig of 2199 bp in length
* 72441 72540: gap of unknown length
* 72541 74292: contig of 1752 bp in length
* 74293 74392: gap of unknown length
* 74393 76503: contig of 2111 bp in length
* 76504 76603: gap of unknown length
* 76604 79013: contig of 2410 bp in length
* 79014 79113: gap of unknown length
* 79114 81669: contig of 2556 bp in length
* 81670 81769: gap of unknown length

Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokoko, S., Ogih, M., Okunou, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojts, A., Rojudoan, L., Rolfe, M., Ruiz, S., Saverly, G.,
Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, Y., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Hoore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 119989)
Worley, K.C.

Direct Submission
Submitted (27-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 119989)
Worley, K.C.

Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20429765.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GVT
Center clone name: CH230-415B14
----- Summary Statistics -----
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 71249 bases at least Q40
Consensus quality: 76269 bases at least Q30
Consensus quality: 79649 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1014: contig of 1014 bp in length
* 1015 1114: gap of unknown length
* 1115 2599: contig of 1485 bp in length
* 2600 2699: gap of unknown length
* 2700 4257: contig of 1558 bp in length
* 4258 4357: gap of unknown length
* 4358 5720: contig of 1363 bp in length
* 5721 5820: gap of unknown length
* 5821 7353: contig of 1533 bp in length
* 7354 7453: gap of unknown length
* 7454 8707: contig of 1254 bp in length
* 8708 8807: gap of unknown length
* 8808 9962: contig of 1155 bp in length
* 9963 10062: gap of unknown length
* 10063 11789: contig of 1727 bp in length
* 11790 11889: gap of unknown length

11890 12898: contig of 1009 bp in length
* 12899 12998: gap of unknown length
* 12999 14485: contig of 1487 bp in length
* 14486 14585: gap of unknown length
* 14586 15746: contig of 1161 bp in length
* 15747 15847: gap of unknown length
* 15848 16902: contig of 1056 bp in length
* 16903 17002: gap of unknown length
* 17003 18294: contig of 1292 bp in length
* 18295 18394: gap of unknown length
* 18395 19731: contig of 1337 bp in length
* 19732 19831: gap of unknown length
* 19832 20840: contig of 1009 bp in length
* 20841 20940: gap of unknown length
* 20941 22880: contig of 1940 bp in length
* 22881 22980: gap of unknown length
* 22981 24470: contig of 1490 bp in length
* 24471 24570: gap of unknown length
* 24571 26235: contig of 1665 bp in length
* 26236 26335: gap of unknown length
* 26336 27539: contig of 1204 bp in length
* 27540 27640: gap of unknown length
* 27641 29666: contig of 2027 bp in length
* 29667 29766: gap of unknown length
* 29767 32144: contig of 2378 bp in length
* 32145 32244: gap of unknown length
* 32245 34551: contig of 2307 bp in length
* 34552 34651: gap of unknown length
* 34652 36879: contig of 2228 bp in length
* 36880 36979: gap of unknown length
* 36980 38695: contig of 1716 bp in length
* 38696 38795: gap of unknown length
* 38796 40176: contig of 1381 bp in length
* 40177 40276: gap of unknown length
* 40277 42504: contig of 2228 bp in length
* 42505 42604: gap of unknown length
* 42605 44375: contig of 1771 bp in length
* 44376 44475: gap of unknown length
* 44476 46073: contig of 1598 bp in length
* 46074 46173: gap of unknown length
* 46174 47658: contig of 1485 bp in length
* 47659 47758: gap of unknown length
* 47759 50216: contig of 2458 bp in length
* 50217 50316: gap of unknown length
* 50317 52271: contig of 1955 bp in length
* 52272 52371: gap of unknown length
* 52372 54575: contig of 2204 bp in length
* 54576 54675: gap of unknown length
* 54676 57590: contig of 2915 bp in length
* 57591 57690: gap of unknown length
* 57691 59891: contig of 2201 bp in length
* 59892 59991: gap of unknown length
* 59992 61347: contig of 1356 bp in length
* 61348 61447: gap of unknown length
* 61448 64459: contig of 3012 bp in length
* 64460 64559: gap of unknown length
* 64560 67954: contig of 3395 bp in length
* 67955 68054: gap of unknown length
* 68055 71223: contig of 3169 bp in length
* 71224 71323: gap of unknown length
* 71324 73860: contig of 2537 bp in length
* 73861 73960: gap of unknown length
* 73961 77898: contig of 3938 bp in length
* 77899 77998: gap of unknown length
* 77999 80364: contig of 2366 bp in length
* 80365 80464: gap of unknown length
* 80465 85763: contig of 5299 bp in length
* 85764 85863: gap of unknown length
* 85864 90800: contig of 4937 bp in length
* 90801 96288: gap of unknown length
* 96289 96388: contig of 5388 bp in length
* 96389 100772: gap of unknown length
* 100773: contig of 4384 bp in length

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, R., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holway, C., Hollins, B., Homst, F., Howard, S., Huber, J., Hui, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Joly, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovich, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lien, C., Liu, J., Liu, W., Loussaged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maneswar, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Plim, E., Fu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojudo, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczek, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 147622)
Worley, K.C.
Direct Submission
Submitted (13-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 147622)
Worley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gl:18846979.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOGF
Center clone name: CH230-240E10
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 99292 bases at least Q40
Consensus quality: 103354 bases at least Q30
Consensus quality: 106749 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1040: contig of 1040 bp in length
* 1041 1140: gap of unknown length
* 1141 2270: contig of 1130 bp in length

2271 2370: gap of unknown length
2371 3422: contig of 1052 bp in length
3423 3522: gap of unknown length
3523 4533: contig of 1011 bp in length
4534 4633: gap of unknown length
4634 6294: contig of 1661 bp in length
6295 6394: gap of unknown length
6395 7900: contig of 1506 bp in length
7901 8000: gap of unknown length
8001 9008: contig of 1008 bp in length
9009 9108: gap of unknown length
9109 10149: contig of 1041 bp in length
10150 10249: gap of unknown length
10250 11661: contig of 1412 bp in length
11662 11761: gap of unknown length
11762 13213: contig of 1452 bp in length
13214 13313: gap of unknown length
13314 14852: contig of 1539 bp in length
14853 14952: gap of unknown length
14953 16042: contig of 1090 bp in length
16043 16142: gap of unknown length
16143 17532: contig of 1390 bp in length
17533 17633: gap of unknown length
17634 18933: contig of 1201 bp in length
18934 18933: gap of unknown length
18934 20169: contig of 1136 bp in length
20170 21285: gap of unknown length
21286 21385: contig of 1116 bp in length
21386 22422: gap of unknown length
22423 22522: gap of unknown length
22523 24767: contig of 2245 bp in length
24768 24867: gap of unknown length
24868 26530: contig of 1663 bp in length
26531 26632: gap of unknown length
26633 28762: contig of 2033 bp in length
28763 29989: contig of 1227 bp in length
29990 30089: gap of unknown length
30090 31439: contig of 1350 bp in length
31440 31549: gap of unknown length
31540 33674: contig of 2135 bp in length
33675 33774: gap of unknown length
33775 35587: contig of 1813 bp in length
35588 35687: gap of unknown length
35689 38101: contig of 2414 bp in length
38102 38201: gap of unknown length
38202 39526: contig of 1325 bp in length
39527 39626: gap of unknown length
39627 41384: contig of 1758 bp in length
41385 41484: gap of unknown length
41485 43477: contig of 1993 bp in length
43478 43577: gap of unknown length
43578 44733: contig of 1156 bp in length
44734 44833: gap of unknown length
44834 46427: contig of 1594 bp in length
46428 46527: gap of unknown length
46529 48229: contig of 1702 bp in length
48230 48339: gap of unknown length
48340 49559: contig of 1230 bp in length
49560 49659: gap of unknown length
49660 50956: contig of 1291 bp in length
50957 51056: gap of unknown length
51057 53501: contig of 2445 bp in length
53502 53601: gap of unknown length
53602 56724: contig of 3123 bp in length
56725 56824: gap of unknown length
56825 59743: contig of 2819 bp in length
59744 59744: gap of unknown length
59745 61622: contig of 1879 bp in length
61623 61722: gap of unknown length
61723 64577: contig of 2855 bp in length
64578 64577: gap of unknown length

* 64678 67244: contig of 2567 bp in length
* 67245 67344: gap of unknown length
* 67345 68774: contig of 1430 bp in length
* 68775 68874: gap of unknown length
* 68875 70933: contig of 2059 bp in length
* 70934 71033: gap of unknown length
* 71034 72774: contig of 1741 bp in length
* 72775 72874: gap of unknown length
* 72875 75760: contig of 2886 bp in length
* 75761 75860: gap of unknown length
* 75861 78701: contig of 2841 bp in length
* 78702 78801: gap of unknown length
* 78802 81489: contig of 2688 bp in length
* 81490 81589: gap of unknown length
* 81590 85093: contig of 3504 bp in length
* 85094 85193: gap of unknown length
* 85194 87862: contig of 2669 bp in length
* 87863 87962: gap of unknown length
* 87963 91250: contig of 3288 bp in length
* 91251 91350: gap of unknown length
* 91351 93945: contig of 2595 bp in length
* 93946 94045: gap of unknown length
* 94046 97369: contig of 3324 bp in length
* 97370 97469: gap of unknown length
* 97470 101504: contig of 4035 bp in length
* 101505 101604: gap of unknown length
* 101605 105465: contig of 3861 bp in length
* 105466 105565: gap of unknown length
* 105566 111708: contig of 6143 bp in length
* 111709 111808: gap of unknown length
* 111809 115961: contig of 4153 bp in length

Query Match 4.0%; Score 62.4; DB 2; Length 147622;
Best Local Similarity 45.3%; Pred. No. 0.00013;

Matches 156; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 1203 GGAAGACAGCAGGAGGCTTAACTATACCTTGTGTTAGAGAAAGACTGTGGCGA 1262
Db 3892 GGGGGGGGGGNNNGGGGGGGGGGNNNNNNNNNNNGGNGGNGGGGGGGGGGG 3951
QY 1263 GGGAGAGAGAGAGAGGAGGCTGAGGAGGAGGCTGAGGAGGAGGAGGAGGAGG 1322
Db 3952 GGG 4011
QY 1323 CAGTGGGGAGGAGGCTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1382
Db 4012 GGGGAGGNNNGGG 4071
QY 1383 GGGGGGGGGGGCTTCTTGCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1442
Db 4072 GGG 4131
QY 1443 CGGCTGAGAGTACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 1502
Db 4132 GGG 4191
QY 1503 GGG 1546
Db 4192 GGG 4235

RESULT 14
AC119304 201650 bp DNA linear HTG 18-JUL-2002
LOCUS AC119304
DEFINITION Rattus norvegicus clone CH230-105f11, *** SEQUENCING IN PROGRESS
ACCESSION AC119304
VERSION AC119304.3 GI:21746367
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE AUTHORS

Rattus.
1 (bases 1 to 201650)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anarlungue,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarel,J., Benton,J., Bimaga,K., Blankenburg,K., Bonin,D.,
Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dethorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H.,
Doutwalte,K.U., Draper,H., Dugan-Kocha,S., Durkin,K.J.,
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsl,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsoson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,T., Li,Z., Lichtarge,O., Liew,C., Liu,J., Liu,W., Louiseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,D., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,Z., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwoudu,G.,
Oraguene,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pyl,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalton,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlaczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 201650)
Worley,K.C.
Direct Submission
Submitted (26-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 201650)
Worley,K.C.
Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20429676.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GUME
Center clone name: CH230-105f11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 134842 bases at least Q40
Consensus quality: 140806 bases at least Q30
Consensus quality: 146272 bases at least Q20

RESULT 15
AC131133
LOCUS
DEFINITION
AC131133
Rattus norvegicus clone CH230-202D8, *** SEQUENCING IN PROGRESS
***, 67 unordered pieces.
AC131133
VERSION
AC131133.1 GI:22296618
HTG: HTG_PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 203343)
REFERENCE
Muzny,D.,Marle., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alshrooks,S., Amla,A., Anguitano,D.,
Anyalebech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Catroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gedregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,U., Kovar,C.,
Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,D.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,L., Loulseged,H., Lozano,R.J., Lu,X., Ma,D.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M., Mcneill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Ming,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Nwokeneme,O., Okwunou,G.,
Olatunpasegon,A., Pal,S., Parks,K., Pasternak,S., Paul,H.,
Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A.,
Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quito,J., Rachlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
Sisson,I., Sitter,C.D., Smaj,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorrelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A.,
Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,R.,
Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlezyk,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausen,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.
Unpublished
2 (bases 1 to 203343)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (17-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GRHP
Center clone name: CH230-202D8

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 142566 bases at least Q40
Consensus quality: 152085 bases at least Q30
Consensus quality: 159708 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraft.html>).
NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1033: contig of 1033 bp in length
1034: gap of unknown length
1133: contig of 1188 bp in length
1134: gap of unknown length
2321: gap of unknown length
2422: contig of 1181 bp in length
3603: gap of unknown length
3702: gap of unknown length
4833: contig of 1131 bp in length
4834: gap of unknown length
4933: contig of 1401 bp in length
6334: gap of unknown length
6335: contig of 1776 bp in length
8210: gap of unknown length
8211: gap of unknown length
8310: gap of unknown length
8311: gap of unknown length
9467: contig of 1157 bp in length
9567: gap of unknown length
10766: contig of 1199 bp in length
10767: gap of unknown length
10866: gap of unknown length
11994: contig of 1128 bp in length
11995: gap of unknown length
12094: gap of unknown length
12095: contig of 1008 bp in length
13102: gap of unknown length
13103: gap of unknown length
13202: gap of unknown length
14497: contig of 1295 bp in length
14498: gap of unknown length
14597: gap of unknown length
14598: contig of 1541 bp in length
16138: gap of unknown length
16139: gap of unknown length
16238: gap of unknown length
17906: contig of 1668 bp in length
17907: gap of unknown length
18007: gap of unknown length
19573: contig of 1567 bp in length
19574: gap of unknown length
19575: gap of unknown length
20704: contig of 1031 bp in length
20705: gap of unknown length
20804: gap of unknown length
22301: contig of 1497 bp in length
22401: gap of unknown length
22402: gap of unknown length
24843: contig of 2442 bp in length
24844: gap of unknown length
24943: gap of unknown length
26418: contig of 1475 bp in length
26419: gap of unknown length
26518: gap of unknown length
28360: contig of 1842 bp in length
28361: gap of unknown length
30423: contig of 1963 bp in length
30424: gap of unknown length
30523: gap of unknown length
30524: gap of unknown length
32559: contig of 2036 bp in length
32560: gap of unknown length
32659: gap of unknown length
34631: contig of 1972 bp in length
34632: gap of unknown length
34731: gap of unknown length
37171: contig of 2440 bp in length
37271: gap of unknown length
37272: gap of unknown length
39423: contig of 2152 bp in length
39424: gap of unknown length
41124: contig of 1601 bp in length
39524

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 04:38:33 ; Search time 366 Seconds
(without alignments)
9555.610 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaacc.....ggtgcgcgtccgcctcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	100.0	1553	24	AAD24913
2	59	3.8	2790	24	AAS94998
3	51.4	3.3	883	22	AAU26741
4	50.4	3.2	3037	22	AAK94933
5	49.4	3.2	7998	22	AAK79404
6	47.6	3.1	407	21	AAK77895
7	47.6	3.1	407	22	AAI28633
8	47.4	3.1	114955	20	AAK53491
9	47	3.0	2691	14	AAQ48043

10	47	3.0	2691	16	AAQ98815	Human beta-IG-H3 (
11	47	3.0	2691	21	AAA77842	CDNA encoding huma
12	47	3.0	2691	22	AAI28580	Colon tumour relat
13	47	3.0	2691	24	ABN95902	Gene #2400 used to
14	47	3.0	2691	24	ABR62672	Colon adenocarcino
15	46.6	3.0	1907	24	ABR99476	Mouse ischemic co
16	46.2	3.0	174424	24	ABL68122	Ovary cancer relat
17	46	3.0	53522	24	AAD30228	Human PKD1 gene.
18	46	3.0	53526	19	AAT94101	Human PKD1 gene.
19	46	3.0	53577	17	AAT18551	Human polycystic k
20	46	3.0	53577	19	AAK94108	Human PKD1 locus b
21	46	3.0	105325	24	ABK94407	DNA encoding endot
22	45.6	2.9	368	22	AAH09431	Human CDNA clone (
23	45.4	2.9	1166	20	AAI19440	M. tuberculosis an
24	45.4	2.9	1166	20	AAI19440	M. tuberculosis re
25	45	2.9	7620	22	ABA07349	Human pancreatic c
26	45	2.9	7620	22	AAK32776	Human genomic DNA
27	45	2.9	114955	20	AAK53491	Human adenosine A1
28	44.4	2.9	1154	24	ABQ68318	Listeria monocytog
29	44.4	2.9	1523	20	AAI17507	Human gene express
30	44.4	2.9	1523	20	AAI17479	Human gene express
31	44.4	2.9	1608	21	AAK22298	Human potassium ch
32	44.2	2.8	1050	22	AAD21685	Mutational hot spo
33	44.2	2.8	2849	22	AAD21684	Human retinitis pi
34	43.6	2.8	357	22	AAI181967	Human polynucleoti
35	43.6	2.8	220895	24	ABK84798	Human CDNA differe
36	43.4	2.8	117213	19	AAV62176	HSV-2 strain SB5 C
37	43.2	2.8	169739	24	ABO88186	Human osteoblast d
38	42.8	2.8	3441	20	AAK80044	Human PRO243 nucle
39	42.8	2.8	3441	21	AAA49552	Human PRO243 CDNA.
40	42.8	2.8	3441	24	ABK40258	CDNA encoding huma
41	42.8	2.8	3441	24	ABK28570	Human DNA359-1207
42	42.6	2.7	600	24	ABO52496	Oligonucleotide fo
43	42.6	2.7	600	24	ABO52497	Oligonucleotide fo
44	42.6	2.7	100301	24	ABO88176	Human osteoblast d
45	42.4	2.7	1382	15	AAQ64890	Human derived adre

ALIGNMENTS

RESULT 1
ID AAD24913 standard; DNA; 1553 BP.
XX AAD24913;
XX
XX 12-MAR-2002 (first entry)
XX
XX Human secreted cell adhesion protein beta IG-H3 promoter DNA.
XX
XX Human: growth inhibitory gene; retinoid; retinoic acid response element;
XX RARE site; therapy; promyelocytic leukaemia; cancer chemoprevention;
XX cytosolic; secreted cell adhesion protein beta IG-H3 promoter; ds.
XX
XX Homo sapiens.
XX
XX WO200192578-A2.
XX
XX 06-DEC-2001.
XX
XX 25-MAY-2001; 2001WO-US17161.
XX
XX 26-MAY-2000; 2000US-207535P.
XX
XX (UNIT) UNIV ILLINOIS FOUND.
XX
XX Roninson IB, Dokmanovic M, Chang B;
XX WPI; 2002-075474/10.
XX
XX Expression construct encoding cellular genes, under control of a
XX promoter regulated by retinoids and cells comprising the construct for
XX PT


```

PF 04-APR-2001; 2001WO-US11128.
XX
PR 05-APR-2000; 2000US-195106P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Shiffman D, Somogyi R, Lawn R, Sellhammer JJ, Porter GJ, Mikita T;
PI Tal J;
XX
XX WPI; 2002-010925/01.
XX
XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development -
XX
XX Claim 1; Page 295-196; 315pp; English.
XX
XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the
XX diagnosis of disorders associated with foam cell development such as
XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as
XX coronary artery disease. The polynucleotide sequences can also be used
XX as PCR primers and probes. The polynucleotides of the invention are also
XX useful in gene therapy. AAS94746-AAS95021 represent the human
XX polynucleotide sequences of the invention which are differentially
XX expressed during foam cell differentiation.
XX
XX Sequence 2790 BP; 708 A; 756 C; 716 G; 610 T; 0 other;
XX
XX Query Match 3.8%; Score 59; DB 24; Length 2790;
XX Best Local Similarity 100.0%; Pred. No. 2.5e-05;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
OY 1495 CTGAGACCGCCCGCTTGCCCGTGGTGCCTGCTGCTGCGCGCTGCCGCTCC 1553
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||
Dd 1 CTGAGACCGCCCGCTTGCCCGTGGTGCCTGCTGCTGCGCGCTGCCGCTCC 59
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3
AAL26741
ID AAL26741 standard; cDNA; 883 BP.
XX
AC AAL26741;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 19198.
XX
XX Human; breast cancer; cell marker; cytostatic; ss.
XX KW Homo sapiens.
XX
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
XX
XX 14-JAN-2000; 2000US-0176077.
XX 14-MAR-2000; 2000US-0189167.
XX 24-MAR-2000; 2000US-0192089.
XX 29-MAR-2000; 2000US-0193480.
XX 15-MAY-2000; 2000US-0205230.
XX 09-JUN-2000; 2000US-0211315.
XX 25-JUL-2000; 2000US-0220534.
XX
XX (MULT-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX

```

XX WP1: 2001-451856/48.

DR New peptide useful as a marker for the diagnosis of breast cancer -

XX

PS Claim 1; Page 3668-3669; 3695pp; English.

XX

CC The invention relates to human breast cancer expressed polynucleotides

CC (AA107544-AA126789) and methods of assessing whether a patient is

CC afflicted with certain cancer by examining the correlation between the

CC expression of certain markers and the cancerous state of breast cells.

CC The polynucleotides and encoded polypeptides are potential markers for

CC detecting, diagnosing, monitoring, characterising treating and

CC potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic

CC activity.

CC

XX

SQ Sequence 883 BP; 175 A; 274 C; 251 G; 183 T; 0 other;

XX

Query Match 3.3%; Score 51.4; DB 22; Length 883;

Best Local Similarity 90.2%; Pred. No. 0.0016;

Matches 55; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1493 CCCGTGAGCCGCCGCTGCGGTGCGTACGTGCTGCGTGGCGTCCCGCTC 1552

DB 13 CCCACGCTCCGCGCGCTGCGGTGCTGCTGCTGCTGCGTGGCGTCCCGCTC 72

OY 1553 C 1553

DB 73 C 73

XX

RESULT 4

ID AAK94933/C AAK94933 standard; cDNA: 3037 BP.

XX

AC AAK94933;

DT 06-NOV-2001 (first entry)

XX

DE Human full-length cDNA, SEQ ID NO: 4176.

XX

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX

OS Homo sapiens.

XX

PN EPI130094-A2.

XX

PD 05-SEP-2001.

XX

PE 07-JUL-2000; 2000EP-0114089.

XX

PR 08-JUL-1999; 99JP-0194486.

XX

PR 11-JAN-2000; 2000JP-0118774.

XX

PR 02-MAY-2000; 2000JP-0183765.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX

WI: 2001-524255/58.

XX

DR P-PSDB; AAM93965.

XX

PT 830 Primers useful for synthesizing full length cDNA clones and their

XX

PS use in genetic manipulation -

XX

XX Claim 8; SEQ ID NO 4176; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

08-NOV-2000: 2000US-0246475.
 PR 08-NOV-2000: 2000US-0246476.
 PR 08-NOV-2000: 2000US-0246477.
 PR 08-NOV-2000: 2000US-0246478.
 PR 08-NOV-2000: 2000US-0246523.
 PR 08-NOV-2000: 2000US-0246524.
 PR 08-NOV-2000: 2000US-0246525.
 PR 08-NOV-2000: 2000US-0246526.
 PR 08-NOV-2000: 2000US-0246527.
 PR 08-NOV-2000: 2000US-0246528.
 PR 08-NOV-2000: 2000US-0246532.
 PR 08-NOV-2000: 2000US-0246539.
 PR 08-NOV-2000: 2000US-0246610.
 PR 08-NOV-2000: 2000US-0246611.
 PR 08-NOV-2000: 2000US-0246613.
 PR 17-NOV-2000: 2000US-0249207.
 PR 17-NOV-2000: 2000US-0249208.
 PR 17-NOV-2000: 2000US-0249209.
 PR 17-NOV-2000: 2000US-0249210.
 PR 17-NOV-2000: 2000US-0249211.
 PR 17-NOV-2000: 2000US-0249212.
 PR 17-NOV-2000: 2000US-0249213.
 PR 17-NOV-2000: 2000US-0249214.
 PR 17-NOV-2000: 2000US-0249215.
 PR 17-NOV-2000: 2000US-0249216.
 PR 17-NOV-2000: 2000US-0249217.
 PR 17-NOV-2000: 2000US-0249218.
 PR 17-NOV-2000: 2000US-0249244.
 PR 17-NOV-2000: 2000US-0249245.
 PR 17-NOV-2000: 2000US-0249264.
 PR 17-NOV-2000: 2000US-0249265.
 PR 17-NOV-2000: 2000US-0249297.
 PR 17-NOV-2000: 2000US-0249299.
 PR 17-NOV-2000: 2000US-0249300.
 PR 01-DEC-2000: 2000US-0250160.
 PR 01-DEC-2000: 2000US-0250391.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 05-DEC-2000: 2000US-0256719.
 PR 06-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251899.
 PR 08-DEC-2000: 2000US-0251990.
 PR 11-DEC-2000: 2000US-0254097.
 PR 05-JAN-2001: 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCT INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 34216; 3071bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially

CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 7998 BP; 1435 A; 2527 C; 2461 G; 1575 T; 0 other;
 Query Match 3.2%; Score 49.4; DB 22; Length 7998;
 Best Local Similarity 57.1%; Pred. No. 0.014;
 Matches 109; Conservative 0; Mismatches 81; Indels 1; Gaps 1;
 QY 1267 GAGAGGAGGAGGAGGCTGCGAGTGAAGGCAAGGCTGGGAAAAGTGAAGCAGCGGACAGT 1326
 DB 1609 GCGGGAGGAGGAGGAGGCGCGGAGTGGGGGCGCGGCG- GGTGGCTGTGCTCCGGAGCGCGG 1551
 QY 1327 GCGGAGGCGGAGGAGGCGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 1386
 DB 1550 TCTGCTGCAAGTGCAGATTTGAGCCAGGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGG 1491
 QY 1387 CCGGCGCTCTTGCAGAGGCGGCGGCGGAGGCTTCCCGCGCGGCGGCGGCGGCGGCGG 1446
 DB 1490 CTCCTTGACTTGGGTGGGCGCGCGCTCTCGGCTCTCCGCTTACGCAAGCGGCTCTACTCTGA 1431
 QY 1447 TCGCAGCTTAC 1457
 DB 1430 GCGGAGCTTAC 1420
 RESULT 6
 ID AAK77895
 XX AAK77895 standard; cDNA: 407 BP.
 AC AAK77895;
 XX
 DT 14-NOV-2000 (first entry)
 XX
 DE cDNA encoding human colon tumour polypeptide. SEQ ID NO:175.
 XX
 KW Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
 KW immunotherapy; diagnosis; progression; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200037643-A2.
 XX
 PD 29-JUN-2000.
 XX
 XX 23-DEC-1999; 99WO-US30909.
 PF 23-DEC-1999; 98US-0221298.
 PR 02-JUL-1999; 99US-0347496.
 PR 22-SEP-1999; 99US-0401064.
 PR 19-NOV-1999; 99US-0444242.
 PR 02-DEC-1999; 99US-0454150.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk J;
 PI Wang T, Yugu J;
 XX
 DR WPI: 2000-442671/38.
 XX
 PT New colon tumor polypeptides used to inhibit the development of cancer,
 PT especially colon cancer, and for diagnosing and monitoring the
 PT progression of the cancer -
 XX
 PS Claim 1; Page 142; 229pp; English.
 XX
 CC Sequences AAK77722-A78199 represent 478 cDNAs encoding proteins or
 CC portions of proteins which are associated with human colon tumours.
 CC The invention also specifically discloses 8 human colon tumour proteins
 CC (AAH1197-B11904). The nucleic acids, the polypeptides they encode, and
 CC antigen presenting cells (APCs, preferably dendritic cells) expressing

PT vasoconstriction
XX
PS Disclosure; Page 37; 120pp; English.
XX
XX The specification describes antisense oligonucleotides (AAK52869-X55274)
CC directed against at least 2 mRNAs selected from target genes, coding and
CC non-coding regions of RNAs corresponding to target genes, gene
CC initiation codons, genomic flanking regions, intron-exon borders, the
CC 5'-end, the 3'-end and the junction between coding and non-coding
CC regions and all segments of RNAs encoding proteins associated with one
CC or more diseases, conditions or mixtures. The antisense oligonucleotides
CC may be derived from sequences AAK55272-74. These multiple target
CC oligonucleotides (specifically AAK5180-271) can be used for the
CC antisense treatment of diseases and conditions. Typical diseases and
CC conditions are those associated with impaired respiration and
CC inflammation, including lung diseases, pulmonary vasoconstriction,
CC inflammation, allergic rhinitis, acute asthma, allergic asthma, impeded
CC respiration, respiratory distress syndrome, pain, cystic fibrosis,
CC pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic
CC obstructive pulmonary disease (COPD), and cancers such as leukemias,
CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,
CC pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,
CC hepatic metastases, as well as all types of cancers which may metastasize
CC or have metastasized to the lungs, including breast and prostate cancer.
XX
SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 3.1%; Score 47.4; DB 20; Length 114955;
Best Local Similarity 38.4%; Pred. No. 0.14;
Matches 156; Conservative 20; Mismatches 229; Indels 1; Gaps 1;

OY 1147 GCATGAGAGACGAGCCGCCGAGACAGTGTCCCTACATGCCAGATGCCCATTTGGAA 1206
DB 108107 GCCTGGCTGCCTTBGGGCCNNNDNNGCGGGGCGCGGCGCTGCTGCTBGGGCS 108166
OY 1207 GCAGCCACAGAGGCTTAAGGACCTATACCTTGGTTTGAGAAAGCTGTGGCAGGGA 1266
DB 108167 NNNDNNGCGGGGCGCGGCGCTGCTGCTBGGGNNNDNNGCGGGGCGCGGCGG 108226
OY 1267 GAGAGGAGGAGGAGGCTGAGTGAAGGAGGCTGGGAAACTGACGAGCGGACAGT 1326
DB 108227 GCCTGGCTGCCTTBGGGNNNDNNGCGGGGCGCGGCGCTGCTGCTBGGGNNNDN 108286
OY 1327 GCGGAGGCGGGTGGTGGCCAGGAGGAGGAGGCGGAGGAGGAGGCGGAGGCGG 1386
DB 108287 GCGGAGGCGGGGCGCGGCTGCTGCTGCTBGGGNNNDNNGCGGGGCGCGGCGG 108346
OY 1387 CCGGCGCTCTTGTGACAGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
DB 108347 GCCTNNNDNNGCGGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108406
OY 1447 TCGACACTTACTTAACTGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1506
DB 108407 GCCTGGCTGCGGNNNDNNGCGGGGCGGCGGCGCTGCTGCTGCTGCTGCTGCTG 108466
OY 1507 GCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552
DB 108467 GC-GGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 108511

RESULT 9
AAQ48043
ID AAQ48043 standard; DNA; 2691 BP.
XX
AC AAQ48043;
XX
XX 07-FEB-1994 (first entry)
DE TGF-beta induced gene betaIG-H3.
XX
XX
KM Transforming growth factor-beta: cell growth; differentiation;
KM Arg-Gly-Asp; RGD; cell surface recognition; communication;
KM Fascilin-I; transmission; intracellular signal; integrin;

KW gene regulation; ss.
XX
XX Homo sapiens.
OS
XX
FH Key
XX CDS
FT Location/Qualifiers
FT 48..209
FT /tag= a
FT /note= "beta IG-H3"
FT 48..89
FT /tag= b
FT 89..90
FT /tag= c
FT /note= "cleavage site"
FT 116..117
FT /tag= d
FT /note= "cleavage site"
FT 452..869
FT /tag= e
FT /number= 1
FT 870..1277
FT /tag= f
FT /number= 2
FT 1278..1658
FT /tag= g
FT /number= 3
FT 1659..2096
FT /tag= h
FT /number= 4
FT 1971..1979
FT /tag= i
FT /note= "RGD sequence"
FT 2672..2677
FT /tag= j

PN EP555989-A.
PD 18-AUG-1993.
XX
XX
XX 04-FEB-1993; 93EP-030809.
PF 05-FEB-1992; 92US-0833835.
PR
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PA
XX
XX Neubauer MG, Purchio AF, Skonler JE;
PI WPI: 1993-260238/33.
PI P-PSDB; AAR40386.
DR
XX
XX TGF-beta induced gene and protein - is designated BIG-H3 contg. 4
PT homologous repeats and possibly involved in cell surface
PT recognition and communication
PS Claim 10; Fig 5; 23pp; English.
XX
XX The betaIG-H3 protein contains 4 homologous repeat regions (at least
CC 16% homology). The protein is possibly a novel cell surface
CC recognition protein (similar to fascilin-I) and may thus have a role
CC in cell-cell communication/transmission of intracellular signals
CC involved in negative growth control. In addition, the protein
CC contains Arg-Gly-asp in the C-terminal which may allow interaction
CC with integrins. The protein may be useful in the study of TGF-beta
CC partic. its role in gene regulation.
XX
SQ Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;

Query Match 3.0%; Score 47; DB 14; Length 2691;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1507 GCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553
DB 1 GCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 47

Sequence	Accession	Species	Gene	Protein	Function	Notes
Sequence 2691 BP, 679 A; 729 C; 695 G; 588 T; 0 other;	AA098815	standard; DNA; 2691 BP.				
	AA098815					
	06-MAR-1996	(first entry)				
	Human beta-IG-H3 (transforming growth factor-beta induced gene-h3).					
	beta-IG-H3; antitumor; ss.					
	Homo sapiens.					
	Key	Location/Qualifiers				
	CDS	48..2099				
		/*tag= a				
	sig_peptide	48..89				
		/*tag= b				
	misc_feature	89..90				
		/*tag= c				
	misc_feature	/note= "Cleaveage site"				
		116..117				
		/*tag= d				
		/note= "Cleaveage site"				
		2669..2677				
	polyA_signal	/*tag= e				
	repeat_region	462..869				
		/*tag= f				
		/note= "REPEAT 1"				
		870..1277				
	repeat_region	/*tag= g				
		/note= "REPEAT 2"				
		1278..1658				
	repeat_region	/*tag= h				
		/note= "REPEAT 3"				
		1659..2693				
	repeat_region	/*tag= i				
		/note= "REPEAT 4"				
	US5444164-A.					
	22-AUG-1995.					
	05-FEB-1992; 92US-0833835.					
	04-MAY-1992; 92US-0878960.					
	05-FEB-1992; 92US-0833835.					
	(BRIM) BRISTOL-MYERS SQUIBB CO.					
	Neubauer MG, Purchio AF, Skonier JE;					
	WPI; 1995-302161/39.					
	P-PSDB; AAR80573.					
	New isolated nucleic acid and protein, beta IG-H3 - which are					
	induced by mammalian cells treated with TGF-beta, used for					
	inhibiting growth of tumour cells					
	Claim 1; Page 11-14; 23pp; English.					
	DNA encoding betaIG-H3 was isolated from a cDNA library prepared					
	from A549 human lung adenocarcinoma cells treated with transforming					
	growth factor-beta. This sequence encodes a beta-IG-H3 protein, the					
	induction, expression and/or secretion of which inhibits the growth					
	colony formation and appearance of tumor cells. The beta-IG-H3 can					
	be used to inhibit the growth of tumor cells.					

Qy	1507	GCATTGCCCGCCGCGCTGCTACGCGCGCGCGCGCTGCTGCGCGTCC	1553
Db	1	GCATTGCCCGCGCGCTGCTACGCGCGCGCGCTGCTGCGCGTCC	47
Query Match 3.0%; Score 47; DB 16; Length 2691;			
Best Local Similarity 100.0%; Pred. No. 0.039;			
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

RESULT 11	
ID	AAA77842
AC	AAA77842 standard; CDNA: 2691 BP.
XX	AAA77842;
XX	14-NOV-2000 (first entry)
DE	CDNA encoding human colon tumour polypeptide, SEQ ID NO:121.
XX	
XX	Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
XX	immunotherapy; diagnosis; progression; beta IG-H3 homologue; ss.
OS	Homo sapiens.
XX	
XX	WO200037643-A2.
XX	
XX	29-JUN-2000.
XX	
XX	23-DEC-1999; 99WO-US30909.
XX	
XX	23-DEC-1998; 98US-0221298.
XX	02-JUL-1999; 99US-0347496.
XX	22-SEP-1999; 99US-0401064.
XX	19-NOV-1999; 99US-0444242.
XX	02-DEC-1999; 99US-0454150.
XX	
XX	(CORI-) CORIXA CORP.
XX	
XX	Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk J;
XX	Wang T, Yugu J;
XX	
XX	MP1: 2000-442671/38.
XX	P-PSDB: AAB11897.
XX	
XX	New colon tumor polypeptides used to inhibit the development of cancer,
XX	especially colon cancer, and for diagnosing and monitoring the
XX	progression of the cancer -
XX	
XX	Claim 29; Page 123; 229pp; English.
XX	
XX	Sequences AAA77722-A78199 represent 478 CDNAs encoding proteins or
XX	portions of proteins which are associated with human colon tumours.
XX	The invention also specifically discloses 8 human colon tumour proteins
XX	(AAB11897-B1904). The nucleic acids, the polypeptides they encode, and
XX	antigen presenting cells (APCs, preferably dendritic cells) expressing
XX	such polypeptides may be used in vaccines that target tumour cells,
XX	especially colon tumour cells, thereby inhibiting the development of
XX	cancer. T-cells specific for the polypeptide expressed by the APC are
XX	used to remove tumour cells from biological samples, especially blood or
XX	fractions thereof. The sample or the isolated T-cells specific for the
XX	polypeptide can then be used to inhibit cancer development. CD4+ and/or
XX	CD8+ T-cells from a patient may be incubated with a polypeptide or
XX	nucleic acid of the invention, or an APC expressing such a polypeptide,
XX	to cause the proliferation of specific T-cells. The T-cells can be
XX	cloned and then administered back to the patient to inhibit cancer
XX	development. Nucleic acids encoding the polypeptides and antibodies
XX	against the polypeptides may be used to determine the expression level
XX	of a tumour protein of the invention, and therefore to determine whether
XX	cancer cells are present. Such diagnostic methods may also be used to
XX	monitor the progression of a cancer by repeating the processes at time
XX	intervals, and comparing the current result to previous results. The
XX	present sequence represents a cDNA encoding a human colon tumour
XX	polypeptide which has homology to beta IG-H3.

XX Sequence 2691 BP; 679 A; 730 C; 695 G; 587 T; 0 other;
XX
Query Match 3.0%; Score 47; DB 21; Length 2691;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Oy 1507 GCCTCCCGTCGCGTACGCAGCCTGCCGTGGCCGCGCCGCCGTC 1553
Db 1 GCTTGCCCGTCGCGTACGTGCTCGGTGGCGGTGTCGCGTC 47

RESULT 12
ID AAI28580 standard; cDNA; 2691 BP.
XX
AAI28580:
AC
XX
AT28580:
DT 12-OCT-2001 (first entry)
DE Colon tumour related full length CDNA sequence SEQ ID NO:121.
XX
KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
KM immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200149716-A2.
PP 12-JUL-2001.
PR
PF 29-DEC-2000; 2000WO-US35596.
XX
PR 30-DEC-1999; 99US-0476296.
PR 10-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX
PA (CORI-) CORIXA CORP.

PJ Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;
XX
DR WPT; 2001-441847/47.

Colon tumor associated proteins and nucleic acids useful for the prevention, diagnosis and treatment of colonic cancer -
PS
PS Claim 25; Page 164-165; 472pp; English.

The present invention describes colon tumour associated proteins (I) and the polynucleotides (II) that encode them. (I) have cytostatic activity. CC (I) and (II) can be used in gene therapy and vaccine production. (I) and CC (II) may be used in the prevention, diagnosis and treatment of diseases CC associated with inappropriate colon tumour associated protein (TCAP) CC expression, such as colonic cancer. For example, (I) and (II) may be CC used to treat disorders associated with decreased expression by CC rectifying mutations or deletions in a patient's genome that affect the CC activity of TCAPs by expressing inactive proteins or to supplement the CC patients own production of them. Additionally, (II) may be used to CC produce the TCAP proteins, by inserting the nucleic acids into a host CC cell culturing the cell to express the protein. (II) and its CC complementary sequences may also be used as DNA probes in diagnostic CC polymerase chain reaction (PCR) and hybridisation assays to detect and CC quantitate the presence of similar nucleic acids in samples, and CC therefore which patients may be in need of restorative therapy. (I) may CC also be used as antigens in the production of antibodies against TCAPs CC and in assays to identify modulators of TCAP expression and activity. Anti-(I) antibodies and antagonists may also be used to down regulate CC TCAP expression and activity. The anti-(I) antibodies may also be used

CC		as diagnostic agents for detecting the presence of TCAPs in samples
CC	(e.g., by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512	
CC	and AAM24494 to AAM24523 represent nucleotide and amino acid sequences	
CC	given in the exemplification of the present invention.	
XX		
SQ	Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;	
	Query Match	3.0%; Score 47; DB 22; Length 2691;
	Best Local Similarity	100.0%; Pred. No. 0.039;
	Matches 47; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1507 GCTTCCCGGTCGTGGTAGCTCGTCGGCGCTGCGCCGCTCC 1553	
Db	1 GCTTCCCGGTCGTGGTAGCTCGTCGGCGCTGCGCCGCTCC 47	
	RESULT 13	
ID	ABN95902	
XX	ABN95902 standard; DNA; 2691 BP.	
AC		
XX	ABN95902;	
XX		
DT	13-AUG-2002 (first entry)	
XX		
DE	Gene #2400 used to diagnose liver cancer.	
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
KV	metastatic liver tumour; cytosolic; expression profile; disease state;	
XX	disease progression; drug toxicity; drug efficacy; drug metabolism.	
OS	Homo sapiens.	
PN	WO200229103-A2.	
PD	11-APR-2002.	
PF	02-OCT-2001; 2001WO-US30589.	
XX		
PR	02-OCT-2000; 2000US-237054P.	
PA	(GENE-) GENE LOGIC INC.	
PI		
PL	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
DR	WPI; 2002-426119/45.	
PT	Diagnosing and detecting the progression of liver cancer,	
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,	
PT	involves detecting the level of expression of two or more genes in a	
XX	liver tissue sample -	
PS	Claim 1; SEQ ID NO 2400; 298pp; English.	
XX		
CC	The invention relates to a novel method for diagnosing and detecting the	
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver	
CC	tumour in a patient, and differentiating metastatic liver cancer from	
CC	hepatocellular carcinoma in a patient, involving detecting the level of	
CC	expression of two or more genes represented in ABN93503-ABN97455 in a	
CC	tissue sample. The method of the invention has hepatotropic, and	
CC	cytostatic activity. The method is useful for diagnosing and detecting	
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic	
CC	liver carcinoma in a patient. The method is useful for identifying	
CC	expression profiles which serve as useful diagnostic markers as well as	
CC	markers that can be used to monitor disease states, disease progression,	
CC	drug toxicity, drug efficacy and drug metabolism.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX	Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;	
	Query Match	3.0%; Score 47; DB 24; Length 2691;
	Best Local Similarity	100.0%; Pred No. 0.039;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1507 GCTGGCCGTCGCTAGCTCGTCGGCGCTCCGCCCTCC 1553
|||||
Db 1 GCTGGCCGTCGCTAGCTCGTCGGCGCTCCGCCCTCC 47

RESULT 14
ABL62672
ID ABL62672 standard; DNA: 2691 BP.

AC ABL62672;
XX
XX
DT 15-MAY-2002 (first entry)
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:1009.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235072P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-23712P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237588P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX

PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 1009; 44pp; English.
XX
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX Sequence 2691 BP; 679 A; 729 C; 695 G; 588 T; 0 other;
SQ

Query Match 3.0%; Score 47; DB 24; Length 2691;
Best Local Similarity 100.0%; Pred. 0.039;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1507 GCTGGCCGTCGCTAGCTCGTCGGCGCTCCGCCCTCC 1553
|||||
Db 1 GCTGGCCGTCGCTAGCTCGTCGGCGCTCCGCCCTCC 47

RESULT 15
ABI9476/c
ID ABI9476 standard; CDNA: 1907 BP.
XX
XX ABI9476;
AC
XX
XX 07-MAR-2002 (first entry)
DT
XX
XX Mouse ischemic condition related CDNA sequence SEQ ID NO:433.
DE
XX
XX Mouse; ischemia; compressive ischaemia; occlusive ischaemia;
KM vasospastic ischaemia; ischemic condition; ischaemic disease; ss.
XX
XX Mus musculus.
OS
XX
XX WO200188188-A2.
PN
XX
XX 22-NOV-2001.
PD
XX
XX 18-MAY-2001; 2001WO-JP04192.
PF
XX
XX 18-MAY-2000; 2000JP-0145977.
PR
XX
XX (UYNT-) UNIV NIHON SCHOOL JURIDICAL PERSON.
PA
XX
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
PI
XX
XX WPI; 2002-034733/04.
DR
XX
XX P-PsDB; ABB57178.
DR
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:26:32 ; Search time 2083 Seconds
(without alignments)
12074.696 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtctcccaacc.....ggtgcgctgctccgctcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_estchum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	7.7	300	9	AU099192
2	92.4	5.9	793	9	AU119221
3	91.8	5.9	433	14	BM843883
4	91	5.9	620	12	BE336857
5	79	5.1	829	12	BE327480
6	79	5.1	842	12	BE775529

7	79	5.1	853	12	BG422226
8	79	5.1	895	9	AL546496
9	70.8	4.6	970	9	AL577405
10	70.2	4.5	987	17	CNS00418
11	67.2	4.3	1021	14	BQ434910
12	66.8	4.3	997	17	CNS005TE
13	66.6	4.3	910	17	CNS006ON
14	65.8	4.2	1101	17	CNS00397
15	65.2	4.2	966	14	BQ672441
16	65	4.2	799	9	AU120568
17	65	4.2	3172	11	BC026352
18	64.8	4.2	1016	14	BQ18616
19	64	4.1	975	12	BG425373
20	64	4.1	1315	13	BM543062
21	63.4	4.1	681	9	AU136140
22	63	4.1	1203	17	CNS015Y4
23	61.8	4.0	1050	17	BQ894806
24	61.6	4.0	1230	13	BM558820
25	60.8	3.9	814	17	AG062282
26	60.6	3.9	928	17	AG130881
27	60.6	3.9	985	14	BQ672571
28	60.6	3.9	1316	13	BM477175
29	60.4	3.9	932	17	CNS00720
30	59.8	3.9	961	17	AG083216
31	59.4	3.8	910	17	CNS006ON
32	59.4	3.8	1176	13	BM544862
33	59.2	3.8	918	12	BC809598
34	59.2	3.8	1364	14	BM810045
35	59	3.8	542	12	BG419864
36	59	3.8	885	12	BG423070
37	59	3.8	892	17	BH133618
38	59	3.8	924	12	BG420398
39	59	3.8	935	17	CNS006XK
40	59	3.8	982	12	BG419111
41	58.8	3.8	732	17	AZ196033
42	58.8	3.8	925	17	CNS0091P
43	58.8	3.8	980	14	BQ228348
44	58.4	3.8	867	12	BG423038
45	58	3.7	925	17	CNS0091P

ALIGNMENTS

RESULT 1
LOCUS AU099192 300 bp mRNA linear EST 05-APR-2001
DEFINITION AU099192 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HRC01126 similar to Human transforming growth factor-beta induced
gene product (BIG3) mRNA, mRNA sequence.
ACCESSION AU099192
VERSION AU099192.1 GI:13550321
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 300)
AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.
TITLE In silico mapping of the 5'-ends of human mRNAs using full-length
enriched and 5'-end enriched cDNA libraries constructed by
Oligo-capping method
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Matuyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

```

FEATURES
  source
    1..300
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="HRC01126"
      /clone_lib="Sugano Homo sapiens cDNA library"
      /note="Differential display comparison of untreated and
      dimethylmitomate treated U937 cells"

BASE COUNT      32 a      117 c      95 g      56 t

ORIGIN
Query Match      7.7%; Score 120; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.7e-19;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1434 CGCTCCCTCCCGCTCCAGCTTACTTAACCTGGCGGCGGAGAGCCCTCCTCCTTC 1493
      |||||||
DB 1 CGCTCCCTCCCGCTCCAGCTTACTTAACCTGGCGGCGGAGAGCCCTCCTCCTTC 60
      |||||||

OY 1494 CCTGAGCCGCGCGCTTCCCGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTCC 1553
      |||||||
DB 61 CCTGAGCCGCGCGCTTCCCGCTGCTGCTAGCTGCTGCTGCTGCTGCTGCTCC 120
      |||||||

RESULT 2
AUI19221 793 bp mRNA linear EST 01-AUG-2002
LOCUS AUI19221 HEMBA1 Homo sapiens cDNA clone HEMBA1005307 5', mRNA
DEFINITION sequence.
ACCESSION AUI19221
VERSION AUI19221
KEYWORDS GI:10934456
SOURCE human.
ORGANISM Homo sapiens
          human.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 793)
          Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Salto,K., Kawai,Y.,
          Yamamoto,T., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
          Isogai,T.
          HRI human cDNA project
          Unpublished (2000)
          Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: genomics@hri.co.jp
          HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
          Research Institute; cDNA library construction: Department of
          Virology, Institute of Medical Science, University of Tokyo, and
          Helix Research Institute.

FEATURES
  source
    1..793
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="HEMBA1005307"
      /clone_lib="HEMBA1"
      /tissue_type="whole embryo, mainly head"
      /dev_stage="embryo; 10 weeks"
      /note="Vector: PME18SFL3"

BASE COUNT      161 a      256 c      216 g      159 t      1 others

ORIGIN
Query Match      5.9%; Score 92.4; DB 9; Length 793;
Best Local Similarity 98.9%; Pred. No. 3.8e-12;
Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1460 AACCTGCGCGCGGCGGAGCGCTCTCCTTCCCTGAGCGCGCGCTTGGCCGCGG 1519
      |||||||
DB 1 AACCTGCGCGCGGCGGAGCGCTCTCCTTCCCTGAGCGCGCGCTTGGCCGCGG 60
      |||||||

```

```

OY 1520 TCCTAGCTCGCTCGCTGCTGCGCGCTCGCTCCGCTCC 1553
      |||||||
DB 61 TCCTAGCTCGCTCGCTGCTGCGCGCTCGCTCCGCTCC 94
      |||||||

RESULT 3
BM843883 433 bp mRNA linear EST 06-MAR-2002
LOCUS BM843883
DEFINITION K-EST0121878 S12SNU216 Homo sapiens cDNA clone S12SNU216-58-G05 5',
          mRNA sequence.
ACCESSION BM843883
VERSION BM843883
KEYWORDS GI:19200292
SOURCE EST.
ORGANISM human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
          1 (bases 1 to 433)
          Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
          Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
          Kim,Y.S.
          21C Frontier Korean EST Project 2001
          Unpublished (2002)
          Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.krdb.re.kr
          Plate: 58 row: 6 column: 05
          High quality sequence stop: 433.

FEATURES
  source
    1..433
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="S12SNU216-58-G05"
      /clone_lib="S12SNU216"
      /sex="F"
      /tissue_type="lymph node"
      /cell_type="Epithelial"
      /cell_line="SNT-216"
      /lab_host="Top10F"
      /note="Organ: Stomach; Vector: PCNS; Site: 1: EcoRI;
      Site: 2: NotI; The poly (A) + RNA was dephosphorylated with
      bacterial alkaline phosphatase (BAP) and then deacapped
      with tabacco acid pyrophosphatase (TAP). The deacapped
      intact mRNA was ligated with DNA-RNA linker including EcoR
      I site by treatment of 14 RNA ligase and the first strand
      cDNA was synthesized from oligo dT-selected mRNA by
      priming with dt-tailed vector. The dt-tailed vector was
      adjusted to have about 60nt. The cDNA vector was
      circularized with E. coli DNA ligase after digestion of
      EcoRI which site is also included in vector. An RNA strand
      converted to a DNA strand by Okayama-Berg method. The
      obtained cDNA vectors were used for transformation of
      competent cells E. coli Top10F' by electroporation method.
      The cDNA libraries constructed by this method are
      full-length enriched cDNA library."

BASE COUNT      76 a      146 c      128 g      82 t      1 others

ORIGIN
Query Match      5.9%; Score 91.8; DB 14; Length 433;
Best Local Similarity 97.9%; Pred. No. 4.3e-12;
Matches 93; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1459 TAACTGCGCGCGGCGGAGCGCTCTCCTTCCCTGAGCGCGCGCTTGGCCGCTG 1518
      |||||||
DB 3 TTAACCTGCGCGCGGCGGAGCGCTCTCCTTCCCTGAGCGCGCGCTTGGCCGCTG 62
      |||||||

OY 1519 GTGCTAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1553
      |||||||
DB 63 GTGCTAGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 97
      |||||||

```

RESULT 4
BE336857
LOCUS
DEFINITION BE336857 620 bp mRNA linear EST 14-JUN-2000
ba97605.y1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:2957889 5
similar to gb:U19932 Mouse (MOUSE);, mRNA sequence.
ACCESSION BE336857
VERSION BE336857.1 GI:9189242
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 620)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: ba97105.x1
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov/image/html/resources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 461.
Location/Qualifiers
1. 620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2957889"
/clone_1lb="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 113 a 207 c 181 g 117 t 2 others
ORIGIN
Query Match 5.9%; Score 91; DB 10; Length 620;
Best Local Similarity 100.0%; Pred. No. 7.8e-12;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1463 CTGCGCCGGGGCGGAGGCGCTCTCACTTCCTGGAGCCGCCGCTGCGCTCG 1522
|||||
DB 1 CTGCGCCGGGGCGGAGGCGCTCTCACTTCCTGGAGCCGCCGCTGCGCTCG 60
QY 1523 CTAGCTCGCTCGGTGGCGCTGCTCCGCTCC 1553
|||||
DB 61 CTAGCTCGCTCGGTGGCGCTGCTCCGCTCC 91
RESULT 5
BG327480
LOCUS
DEFINITION BG327480 829 bp mRNA linear EST 27-FEB-2001
602426462F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4564383 5',
mRNA sequence.
ACCESSION BG327480
VERSION BG327480.1 GI:13133918
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 829)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
High quality sequence stop: 704.
Location/Qualifiers
1. 829
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4564383"
/clone_1lb="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 172 a 268 c 233 g 156 t
ORIGIN
Query Match 5.1%; Score 79; DB 12; Length 829;
Best Local Similarity 100.0%; Pred. No. 8.3e-09;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1475 GCGGAGCGCTCTCACTTCCTGGAGCCGCCGCTGCGCTGCGCTCG 1534
|||||
DB 1 GCGGAGCGCTCTCACTTCCTGGAGCCGCCGCTGCGCTGCGCTCG 60
QY 1535 GTGCGCTCGTCCGCTCC 1553
|||||
DB 61 GTGCGCTCGTCCGCTCC 79
RESULT 6
BG775529
LOCUS
DEFINITION BG775529 842 bp mRNA linear EST 15-MAY-2001
602650440F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761465 5',
mRNA sequence.
ACCESSION BG775529
VERSION BG775529.1 GI:14045846
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 842)
AUTHORS NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1CM1613 row: p column: 10
High quality sequence stop: 787.
Location/Qualifiers

FEATURES

source

1..842

/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone_lib="NIH_MGC_40"
/issue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTR; Site:1: XhoI;

Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(g). Library constructed by
Liang Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

BASE COUNT 179 a 269 c 230 g 164 t

ORIGIN

Query Match 5.1%; Score 79; DB 12; Length 842;
Best Local Similarity 100.0%; Pred. No. 8.3e-09;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 GCGGAGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 1534

DB 1 GCGGAGCGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 60

QY 1535 GTGCGCGTCTCCCGCTCC 1553

DB 61 GTGCGCGTCTCCCGCTCC 79

RESULT 7 853 bp mRNA linear EST 14-MAR-2001
BG422226 60244687F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4585263 5',

LOCUS BG422226 mRNA sequence.

ACCESSION BG422226.1 GI:13328732

VERSION EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone_lib="NIH_MGC_14"
/issue_type="renal cell adenocarcinoma"

/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTR; Site:1: XhoI; Site:2:

EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(g). Size-selected >500bp for average

Insert size 1.8kb. Library constructed by Liang Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 179 a 269 c 238 g 167 t

ORIGIN

Query Match 5.1%; Score 79; DB 12; Length 853;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 GCGGAGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 1534

DB 1 GCGGAGCGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 60

QY 1535 GTGCGCGTCTCCCGCTCC 1553

DB 61 GTGCGCGTCTCCCGCTCC 79

RESULT 8 895 bp mRNA linear EST 16-FEB-2001
AL546496 AL546496 LTI_NFL006_P12 Homo sapiens cDNA clone CSOD1030YC01 5

LOCUS AL546496 prime, mRNA sequence.

ACCESSION AL546496.1 GI:12879668

VERSION EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2001)
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

LOCATION/Qualifiers

1..895

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="LTI_NFL006_P12"

/issue_type="placenta"

/note="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 189 a 285 c 245 g 174 t 2 others

ORIGIN

Query Match 5.1%; Score 79; DB 9; Length 895;
Best Local Similarity 100.0%; Pred. No. 8.5e-09;

Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 GCGGAGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 1534

DB 1 GCGGAGCGCGCTCTACTTCCCTGGAGCCGCCGCTTCCCGTGGCTAGCTCGCTCG 60

QY 1535 GTGCGCGTCTCCCGCTCC 1553

DB 61 GTGCGCGTCTCCCGCTCC 79

[illegible]

COMMENT	Determination of this BAC-end sequence was carried out as part of collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .					
FEATURES	Location/Qualifiers					
SOURCE	1..987 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACR03C16" /clone_1id="RPCL-98" /note="end : TET3"					
BASE COUNT	238 a 162 c 17 g 177 t 393 others					
ORIGIN						
Query Match	4.5%; Score 70.2; DB 17; Length 987;					
Best Local Similarity	24.3%; Pred. No. 1.4e+06;					
Matches	96; Conservative 128; Mismatches 171; Indels 0; Gaps 0;					
OY	19	CCAGCAGCTACGTCGGGAAGCGTGCATGTGCCTCATTTCTGAACATGGCACAT	78			
Db	521	YYYYCYCCCTCYCCGCCCCGCCCYCCYCCYCCYCCCTCCCTCCYCTTTTTYY	580			
OY	79	CACCTGCTCAATGTCACAGACACAACACCCTCAATAATTCCTGACTCCCTTCGCC	138			
Db	581	CTYYCYTYYYYTCYTCTCTYTYTYTYCYCYCTYCTCTYCTYTYTYTTTTCY	640			
OY	139	TTCCTCCGTATACAGACTCCAAATTGTAATGAGACTAATACCTCTACACCCTC	198			
Db	641	CYCCCTCTCTCYCCYCCYCTYTYTYTYCTGYTYTYCTCYCCYCCYCCYCCCTCC	700			
OY	199	TTGCCACGCTCCGCATCTGCCTCAGCACCATATTTCAAGCTTCGCCATGCTCCT	258			
Db	701	TCT	760			
OY	259	TCTGCTTACCTGTTCT	318			
Db	761	YCYTY	820			
OY	319	CCTCCACCCATACCCGCTAGAGCTTATAGCTGAGTGGGCCCTGCTACCTCTCTCC	378			
Db	821	YCYYCYCTY	880			
OY	379	TTCCTCCCTACCTGATCCACCTCCCTGCTGCT	413			
Db	881	YTCCT	915			
RESULT 11						
LOCUS	BO434910 1021 bp mRNA linear EST 24-MAY-2002					
DEFINITION	AGNCNOCURT_7909708 NIH_MGC_71 Homo sapiens cdna clone IMAGE:6156444					
ACCESSION	BO434910					
VERSION	BO434910.1 GI:21173966					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1021) NIH-MGC http://mgc.ncl.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC)					
AUTHORS						
TITLE						

Source

1. .910

202. a 63 c 112 g

```

best local similarity  22.7%;  pred.no. 1e-05;
Matches    71;  Conservative 125;  Mismatches 117;  Indels    0;  Gaps    0

```

71; Conservative 125; Mismatches 117; Indels 0; Gaps 0;

20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050

[illegible]

```

550 :::: :: | :: | : :::: | : |
      :::: :: | :: | : :::: | : |

```

[illegible]

Db 619 SSSTSGCGYSSSGSBSSSCSSTBTSSSTTBTBTBKBTSTYSTSTSTTTSTTSTBSBTSYG 678

1479 AGCGCTCTCACTTCCCTGGAGCCGCCGCTTGCCCGTGGTGGCTAGCTCGCTGGTGC 153

Db 679 SBBKSTSSSTSSSTSSSCTSTKBYSTSTBSYBTBTBTSTSTBTBTST 738

QY 1539 GCGTCGTCGCCGCT 1551

DB / 39 BY:BBJ:IBY:IKS: / 51

RESULT 14

LOCUS	1101 bp	DNA	linear	GSS 03-JUN
CNS00397				

BACR08A10 of RPCI-98 library from *Drosophila melanogaster* (fru

ACCESSION AL063912

KEYWORDS GSS.
Drosophila melanogaster

ORGANISM
Drosophila melanogaster
Eukaryota: Metazoa: Hexanoda: Insecta: Pterygota:

neoptera; Eucopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidae; Drosophilidae; Drosophila.

AUTHORS Genoscope.

FEATURES

Location/Qualifiers

111 a 286 c 221 g

Best Local Similarity	32.9%	Pred. No.	1.7e-05;
Matches	140;	Conservative	86;
		Mismatches	199;
		Indels	0;
		Gaps	0

Matches 140; Conservative 86; Mismatches 199; Indels 0; Gaps 0

1050 TGGGCATGCTTCTGGGAATCAGCAGTCCCTCCCTCAGTGTTC 1100

Db 10/5 BRBGRKGGKKYKGRKGTGGSGBGCTYGTYYBKRCGTCCCKTYGYGCSTCGKBByGtG 101

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

A95 TCTGCGCCGCAGBACCATCACCACCAATCACGCCAACAACAAACAAA :
| | :: | : ||::||| || || :

b6

1290 GAGGGCAAGGCTGGGAAACTGAGGCACGGGCACAGTGGGGGAGCGGGTGGGTGCTTACG 1340

1350 GCAGCCAGGGCGCACGGGTTGGGAGGCCCAAGGGCCCGCCCTCCTTGCAAGGGCGG 1400

Db 775 GGGGGGGGGGGGGGGGGGGGGRGGGGSGRASCAGAGACCRSGSVCg 716

1410 CCAGCTTCCCCGCTGGCGTCCGCTCCGCTCGCAGTTACTTAACCTGGCCC 1460

DB /15 GGGGGGGGCCSCCGGGGCCGACCCCVCMCCGSGCGGSCGCGCCCMCCSSGCC 656

27 14/0 55555 14/4

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 15

LOCUS	BQ672441	966 bp	mRNA	linear	EST 15-JUL
-------	----------	--------	------	--------	------------

5', mRNA sequence.
 B0672441
 ACCESSION

NOTATION
DXX
T:TT442/00
C/759/17:TS
GT
17032/C

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 966)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LNCM2460 row: 1 column: 08
High quality sequence start: 2
High quality sequence stop: 434.
Location/Qualifiers

FEATURES

source
1. 966
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6276247"
/clone_lib="NIH_MGC_102"
/issue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 128 a 271 c 427 g 88 t 52 others
ORIGIN

Query Match 4.2%; Score 65.2; DB 14; Length 966;
Best Local Similarity 54.0%; Pred. No. 2, 3e-05;
Matches 127; Conservative 0; Mismatches 108; Indels 0; Caps 0;

QY 1241 GGTTCAGGAACTGCTGCGAGGAGAGAGAGGAGGCTGCACTGAGGCAAGG 1300
|| || || || || || || || || || || || || || || || || || || ||
Db 565 GGG 624
QY 1301 CTGGGAAACTGACACGGGACACAGTCGGGAGCGGCTGCCAGGCGACGAGG 1360
|| || || || || || || || || || || || || || || || || || || ||
Db 625 GGNCGGGGAAAGTGNGGGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 684
QY 1361 CGACAGGCTTGGAGGCGCGAGCGCGCCCTCTTGCACGGGCGGCGCACACTTCC 1420
|| || || || || || || || || || || || || || || || || || || ||
Db 685 GGGCGGGGGGGGGGAGCTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 744
QY 1421 CGCCCTGCGCTCCCTCCCTCCCTGCGAGCTTACTTAAGCTGCGCGGCGG 1475
|| || || || || || || || || || || || || || || || || || || ||
Db 745 CCCCGCGCGCGCGCGCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 799

Search completed: February 23, 2003, 08:21:29
Job time : 2104 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:33:27 ; Search time 87 Seconds
(without alignments)
5474.356 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaacc.....ggtcgcgtcgtccgcgtcc 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents.NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96.6	6.2	7218	1	US-08-232-463-14
2	47.6	3.1	7218	1	US-08-232-463-14
3	47.6	3.0	2691	1	US-07-878-960-1
4	46	3.0	53526	3	US-08-658-136-2
5	46	2.9	53577	3	US-08-658-136-1
6	45.4	2.9	1166	4	US-09-072-596-323
7	41.6	2.7	4897	6	5196516-7
8	41.4	2.7	2269	4	US-09-394-645-1
9	41.4	2.7	2269	4	US-09-243-560B-1
10	40.6	2.6	289	4	US-09-007-005-17
11	40.6	2.6	289	4	US-09-244-796-17
12	40	2.6	8438	1	US-07-945-283-1
13	39.6	2.5	3480	4	US-09-226-012-1
14	39.6	2.5	3950	4	US-09-226-012-3
15	39.2	2.5	320	4	US-09-165-264-7
16	39.2	2.5	2353	5	PCT-US92-06840-1
17	38.8	2.5	1719	4	US-09-330-740A-9
18	38.8	2.5	152331	3	US-09-128-155-16
19	38.6	2.5	351	4	US-09-220-528-7
20	38.6	2.5	351	4	US-09-220-528-10
21	38.6	2.5	423	4	US-09-220-528-8
22	38.6	2.5	423	4	US-09-220-528-11
23	38.6	2.5	546	4	US-09-220-528-42
24	38.6	2.5	663	4	US-09-220-528-24
25	38.6	2.5	663	4	US-09-220-528-25
26	38.6	2.5	663	4	US-09-220-528-44
27	38.6	2.5	696	4	US-09-220-528-1

C 28	38.6	2.5	696	4	US-09-220-528-2	Sequence 2, Appl
C 29	38.6	2.5	714	4	US-09-220-528-30	Sequence 30, Appl
C 30	38.6	2.5	714	4	US-09-220-528-31	Sequence 31, Appl
C 31	38.6	2.5	1652	4	US-09-220-528-68	Sequence 68, Appl
C 32	38.6	2.5	1652	4	US-09-220-528-69	Sequence 69, Appl
C 33	38.6	2.5	1763	6	5198542-1	Patent No. 5198542
C 34	38.4	2.5	1464	1	US-08-243-542-5	Sequence 5, Appl
C 35	38.4	2.5	1464	1	US-08-477-407-5	Sequence 5, Appl
C 36	38.4	2.5	1464	1	US-08-484-355-5	Sequence 5, Appl
C 37	38.4	2.5	2913	1	US-08-243-542-7	Sequence 7, Appl
C 38	38.4	2.5	2913	1	US-08-477-407-7	Sequence 7, Appl
C 39	38.4	2.5	2913	1	US-08-484-355-7	Sequence 7, Appl
C 40	38.4	2.5	2923	1	US-08-243-542-6	Sequence 6, Appl
C 41	38.4	2.5	2923	1	US-08-477-407-6	Sequence 6, Appl
C 42	38.4	2.5	2923	1	US-08-484-355-6	Sequence 6, Appl
C 43	38.4	2.5	3183	1	US-08-243-542-8	Sequence 8, Appl
C 44	38.4	2.5	3183	1	US-08-477-407-8	Sequence 8, Appl
C 45	38.4	2.5	3183	1	US-08-484-355-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMUG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14
Query Match 6.2%; Score 96.6; DB 1; Length 7218;

Best Local Similarity 7.08; Pred. No. 2.2e-16;
Matches 30; Conservative 254; Mismatches 143; Indels 0; Gaps 0;

```
OY 58 TCATTCGACATGACGACATGCTGCAATGTCACAGACCAACACCTACCAATAT 117
   | |||| | | |||| | | | | | | | | | | | | | | | | | | | | | |
Db 1028 TAATTCGACATGCTGCTGACAGCTGAGGAGCTTGCATTTTTTTTTTTTTTTT 1087
OY 118 CCTTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1088 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1147
OY 178 TACCTCTACACCCCTACATTTGCCAGCCTTCCCATCTCTCTCTCTCTCTCTCT 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1148 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1207
OY 238 TCAAGCTCTCCAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1208 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1267
OY 298 TGACACTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1268 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1327
OY 358 GGGCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1328 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1387
OY 418 CACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1388 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1447
OY 478 AACTTTC 484
   | | | |
Db 1448 CTAATTC 1454
```

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109
TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 3.18; Score 47.6; DB 1; Length 7218;
Best Local Similarity 8.38; Pred. No. 0.0041;
Matches 17; Conservative 120; Mismatches 69; Indels 0; Gaps 0;

```
OY 1203 GGAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1238 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1179
OY 1263 GGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1178 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1119
OY 1323 CAGTGGGAGCGGGTGTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1118 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1059
OY 1383 GGGCGCGCCCTCTCTGACAGGCGG 1408
   | | | | | | | | | |
Db 1058 CTCCTGACCTGACGACGAGCTCGG 1033
```

RESULT 3
US-07-878-960-1
; Sequence 1, Application US/07878960
; Patent No. 544164
; GENERAL INFORMATION:
; APPLICANT: Purchio, Anthony F.
; APPLICANT: Skonier, John
; APPLICANT: Neubauer, Michael G.
; TITLE OF INVENTION: TGF-BETA INDUCED GENE AND PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bristol-Myers Squibb Company
; STREET: 3005 First Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/878,960
; FILING DATE: 05-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/833,835
; FILING DATE: 05-FEB-1992
; NAME: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Sorrentino, Joseph M.
; REGISTRATION NUMBER: 32,598
; REFERENCE/DOCKET NUMBER: ON0092-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206/728-4800
; TELEFAX: 206/727-3601
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2691 base pairs

```

?      type: nucleic acid
?      STRANDNESS: double
?      MOLECULE TYPE: linear
?      MOLECULE TYPE: CDNA
?      HYPOTHETICAL: NO
?      ORIGINAL SOURCE:
?      ORGANISM: Homo sapiens
?      TISSUE TYPE: LUNG
?      CELL TYPE: ADENOCARCINOMA
?      CELL LINE: A549
US-07-878-960-1

```

Query Match	3.08;	Score 47;	DB 1;	Length 2691;
Best Local Similarity	100.08;	Pred. No.	0.0039;	
Matches 47: Conservative	0:	Mismatches	0:	Indels

QY 1507 GCTTGCCCGTCGGTGCCTAGCTCGCTCGGTGCGCGCTCGTCCC GGCTCC 1553
Dd 1 GCTTGCCCGTCGGTGCCTAGCTCGCTCGGTGCGCGCTCGTCCC GGCTCC 47

RESULT 4
US-08-658-136-2
; Sequence 2, Application US/08658136

GENERAL INFORMATION:
 APPLICANT: KLINGER, KATHERINE W
 APPLICANT: LANDES, GREGORY M
 APPLICANT: BURN, TIMOTHY C
 APPLICANT: CONNORS, TIMOTHY D
 APPLICANT: DACKOWSKI, WILLIAM
 APPLICANT: GERMINO, GREGORY
 APPLICANT: QIAN, FENG
 TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENES
 NUMBER OF SEQUENCES: 58
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENZYME CORPORATION
 STREET: ONE MOUNTAIN ROAD
 CITY: FRAMINGHAM
 STATE: MASSACHUSETTS
 COUNTRY: USA
 ZIP: 01701
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/658,136

```

? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: LASEN, ELIZABETH
? REGISTRATION NUMBER: 31,845
? REFERENCE/DOCKET NUMBER: GENA-17,8
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 508-872-8400
? TELEFAX: 508-872-5415
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 53526 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
IS-O8-658-136-2

```

Query Match 3.08; Score 46; DB 3; Length 53526;

Matches 144; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

OY 126 CTCTTTTCCTCCCTTCTCCTGTATACAGACTCCAAATTCTAATGAGACTATTAECTCCT 185
 |||| | |||| | | | |||| | | | | |
 |||| | |||| | | | |||| | | | | |

[illegible]

Db 34828 CTGCCCTCTCTCTCTCTCACCTCCCTTCTCCGCTCTTCTTCTCTC 34877

RESULT 5
US-08-658-136-1
; Sequence 1, Application US/08658136

1 GENERAL INFORMATION:
 2 APPLICANT: KLINGER, KATHERINE W
 3 APPLICANT: LANDES, GREGORY M
 4 APPLICANT: BURN, TIMOTHY C
 5 APPLICANT: CONNORS, TIMOTHY D
 6 APPLICANT: DACKOWSKI, WILLIAM
 7 APPLICANT: GERMINO, GREGORY
 8 APPLICANT: OLAN, FENG
 9 TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
 10 NUMBER OF SEQUENCES: 58
 11 CORRESPONDENCE ADDRESS:
 12 ADDRESSEE: GENZYME CORPORATION
 13 STREET: ONE MOUNTAIN ROAD
 14 CITY: FRAMINGHAM
 15 STATE: MASSACHUSETTS
 16 COUNTRY: USA
 17 ZIP: 01701
 18 COMPUTER READABLE FORM:
 19 MEDIUM TYPE: Floppy disk
 20 COMPUTER: IBM PC compatible
 21 OPERATING SYSTEM: PC-DOS/MS-DOS
 22 SOFTWARE: PatentIn Release #1.0, Version #1.25
 23 CURRENT APPLICATION DATA:
 24 APPLICATION NUMBER: US/08/658,136

```

? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: LASSEN, ELIZABETH
? REGISTRATION NUMBER: 31,845
? REFERENCE/DOCKET NUMBER: GEN4-17, 8
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 508-872-8400
? TELEFAX: 508-872-5415
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 53577 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? OS-08-658-136-1

```

Query Match	Score	DB	Length
3.08;	46;	3;	53577;

Matches 144; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

126 CTCCTTCTCCCTTCTCCCTGTATACAGACTCCCAATCTATTGAGACTATTACCTCCT 185


```

Db      1954 CTCGGCGGCGGGGTCTTCCTCCCGCTTCCCGCAGTTCAACAGGGGGGTCTTGACCAACTGCC 2013
Oy      1370 TGGGAGGGGCCAGGGGGCCCCGCCTTCCTTGACAGGGGCGGCCAGCTTCCCGCCCTGG 1429
Db      2014 TCACATCCCGCCCGAGGGGGGAGGCTCAAAAGCCCCTTGCCCGCCCTTCCTCCCGCCCTGA 2073
Oy      1430 GGTCCGCTCCCTCCCGCTCGCAGCTTACTTAACTGAGCCGGG 1472
Db      2074 ACCCGGGCCCGCGGGGCGGGCGCGCTTCTGCGCCCGG 2116

RESULT 9
US-09-243-560B-1
; Sequence 1, Application US/09243560B
; Patent No. 6395882
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Sassetti, Christopher M.
; TITLE OF INVENTION: No. 6395882el Selectin ligands
; FILE REFERENCE: UCAL097U51
; CURRENT APPLICATION NUMBER: US/09/243, 560B
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: US 60/074,389
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-243-560B-1

Query Match          2.7%; Score 41.4; DB 4; Length 2269;
Best Local Similarity 53.4%; Pred. No. 0.12;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Oy      1310 CTGACACAGGGGCACAGTGTGCGGAGCGGGTGGTGCCCAAGGCGACGCGCGCACGGGT 1369
Db      1954 CTCGGCGGCGGGGTCTTCCTCCCGCTTCCCGCAGTTCAACAGGGGGCTTGACCAACTGCC 2013
Oy      1370 TGGGAGGGGCCAGGGGGCCCCGCCTTCCTTGACAGGGGCGGCCAGCTTCCCGCCCTGG 1429
Db      2014 TCACATCCCGCCCGAGGGGGGAGGCTCAAAAGCCCCTTGCGCCCGCTTCCTCCCGCCCTGA 2073
Oy      1430 GGTCCGCTCCCTCCCGCTCGCAGCTTACTTAACTGAGCCGGG 1472
Db      2074 ACCCGGGCCCGCGGGGCGGGCGCGCTTCTGCGCCCGG 2116

RESULT 10
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihue
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence

```

[illegible]

Oy	166	TATTAGCTATTACCTCTACACCCCTACATTTGCCACACCTCCCATCTGCTC	225
Db	145	YNYNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYS	86
Oy	226	TACACCATTAAGTTCAGCTCTCCCATGGTCCCTTCGGTTACCTGTTTC	274
Db	85	YNYNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYSYNNYS	37
RESULT 12			
US-07-945-283-1			
Sequence 1, Application US/07945283			
Patent No. 5352596			
GENERAL INFORMATION:			
APPLICANT: Cheung, Andrew K.			
APPLICANT: Wesley, Ronald D.			
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants			
TITLE OF INVENTION: Involving the EP0 and LTR Genes			
NUMBER OF SEQUENCES: 7			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Curtiss P. Ribando			
STREET: 1815 No. 5352596th University Street			
CITY: Peoria			
STATE: IL			
COUNTRY: USA			
ZIP: 61604			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/945,283			
FILING DATE: 19920911			
CLASSIFICATION: 424			
ATTORNEY/AGENT INFORMATION:			
NAME: Ribando, Curtiss P.			
REGISTRATION NUMBER: 27976			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 309-685-4011 ext. 513			
TELEFAX: 309-685-4128			
INFORMATION FOR SEQ ID NO: 1:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 8438 base pairs			
TYPE: NUCLEIC ACID			
STRANDEDNESS: double			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
HYPOTHETICAL: NO			
ANTI-SENSE: NO			
ORIGINAL SOURCE:			
ORGANISM: Pseudorabies virus			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 622..6495			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1099, "g")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1267, "c")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1381, "c")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(1566, "c")			
FEATURE:			
NAME/KEY: variation			
LOCATION: replace(7010, "g")			
US-07-945-283-1			

Query Match	2.6%	Score 40	DB 1	Length 8438
Best Local Similarity	51.7%	Pred. No. 0.5		
Matches	91	Conservative	0	Mismatches 85; Indels 0; Gaps 0
QY	1254	CTGTGGCGAGGAGAGAGAGGAGGAGGCTGTGCACTGTGAGGGCAAGGGCTGGGAAACTGA	1313	
Db	6217	CCCCGGGGCGCCGAGAGGGGCGGGGAGCCGAGAGAGCCGAGAGAGCCGAGAGAGAGGA	6276	
QY	1314	GCAAGGGCACAGTGTGGGAGACGGGTGTGGTCCCAAGGCGACAGGGGCGACAGGGTTGGG	1373	
Db	6277	GGCGGGGGAAGCTGTGCGAGAGGGGACAGAGGCGCCGGGGGCGCGCGGGGGGCCCGCGGCTCTTGG	6336	
QY	1374	AGGCGCCAGGCGGGCGCCGCTCTTTCACAGGGCGCGCCACACTTTCGCCGCGCCCTMG	1429	
Db	6337	CGCTGCACAGTGTGTCCGGGGGTCCGGGCGCCCGGGGACCCCGAGCGCGGCCCGGAGACCCGG	6392	

```

RESULT 13
US-09-226-012-1/c
Sequence 1, Application US/09226012
Patent No. 6207383
GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
APPLICANT: Splawski, Igor
TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
FILE REFERENCE: 2323-136
CURRENT APPLICATION NUMBER: US/09/226,012
CURRENT FILING DATE: 1999-01-06
EARLIER APPLICATION NUMBER: 09/122,847
EARLIER FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 116
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3480
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(3477)
US-09-226-012-1

```

Query Match	2.5%	Score 39.6	DB 4	Length 3480
Best Local Similarity	51.7%	Pred. No. 0.44	Mismatches 84	Indels 0
Matches 90	Conservative	0	Gaps	0
QY 1280	GCGTGCAGTGAAGGGCGCTGGGAAAAC	TGAGACACAGGGGACACATGTGCGGAGCGGGT	1339	
Db 2900	GCGTCTCCGGGGGCGCTGGGGCTGGAGAAAGG	CACACAGCGGAGGGGGCTGAGACTGGCG	2841	
QY 1340	GGTCCAGGGGACGACGAGGGCGCACAGGGTT	TGGAGGCGCCAGGCGCCCGCCCTCTCTGG	1399	
Db 2840	CTTGGGCCCTCATCTCACTGCTCTCAGGGG	CGTGGGCGGCACATGTGAGCGGGGCTCTCCCC	2781	
QY 1400	CAGGGGCGGGCCACGCTTCCCGGGCCCGGG	GGGTCCGCTCCCTCCGCTCGCACG	1453	
Db 2780	CAGGGCCCCCGGGCGGGCCCGGCTACTCG	GGCTCTGGCCCCCGGCGGCCGGC	2727	

```

: RESULT 14
: US-09-226-012-3/c
: Sequence 3, Application US/09226012
: Patent No. 6207383
: GENERAL INFORMATION:
: APPLICANT: Keating, Mark T.
: APPLICANT: Splawski, Igor
: TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
: TITLE OF INVENTION: SYNPROME GENE
: FILE REFERENCE: 2323-136
: CURRENT APPLICATION NUMBER: US/09/226,012
: CURRENT FILING DATE: 1999-01-06
: EARLIER APPLICATION NUMBER: 09/122,847
: EARLIER FILING DATE: 1998-07-27

```


GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 23, 2003, 06:41:08 ; Search time 142 Seconds

(without alignments)
6142.437 Million cell updates/sec

Title: US-09-865-879-2

Perfect score: 1553
Sequence: 1 gactcagggtgtcccaaac.....ggtagcgctgcgcgcctcc 1553Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44218 seqs, 280819700 residues

Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	3.3	2782	12	US-10-044-090-647
2	47.6	3.1	407	9	US-10-025-380-175
3	47.6	3.1	407	10	US-09-922-217-175
4	47.6	3.1	407	10	US-09-833-263-175
5	47	3.0	2691	9	US-10-025-380-121
6	47	3.0	2691	10	US-09-922-217-121
7	47	3.0	2691	10	US-09-833-263-121
8	47	3.0	2691	10	US-09-880-107-2399
9	46.2	3.0	174424	10	US-09-967-768A-114
10	46	3.0	53522	9	US-09-904-968A-1
11	42.8	2.8	3441	9	US-09-944-413-6
12	42.8	2.8	3441	9	US-09-944-403-6
13	42.8	2.8	3441	9	US-09-944-896-6
14	42.8	2.8	3441	9	US-09-944-944-6
15	42.8	2.8	3441	9	US-09-944-907-6
16	42.8	2.8	3441	9	US-09-944-929-6
17	42.8	2.8	3441	10	US-09-866-028-6
18	42.8	2.8	3441	10	US-09-944-449-6
19	42.8	2.8	3441	10	US-09-944-457-6

20	42.8	2.8	3441	10	US-09-945-587-6	Sequence 6, Appl1
21	42.8	2.8	3441	10	US-09-945-015-6	Sequence 6, Appl1
22	42.8	2.8	3441	10	US-09-944-396-6	Sequence 6, Appl1
23	42.8	2.8	3441	10	US-09-944-097-6	Sequence 6, Appl1
24	42.8	2.8	3441	10	US-09-944-432-6	Sequence 6, Appl1
25	42.8	2.8	3441	10	US-09-943-762-6	Sequence 6, Appl1
26	42.8	2.8	3441	10	US-09-944-654-6	Sequence 6, Appl1
27	42.8	2.8	3441	10	US-09-943-851A-6	Sequence 6, Appl1
28	42.4	2.7	5025	10	US-09-960-253-176	Sequence 176, App
29	42.4	2.7	42999	9	US-09-799-462A-17	Sequence 17, Appl
30	42.4	2.7	42999	9	US-10-125-767-17	Sequence 17, Appl
31	42.4	2.7	42999	9	US-09-836-911A-17	Sequence 17, Appl
32	41.8	2.7	312	10	US-09-864-761-18121	Sequence 18121, A
33	41.8	2.7	1668	10	US-09-948-018-1	Sequence 3, Appl1
34	41.8	2.7	2223	9	US-09-999-248-3	Sequence 1, Appl1
35	41.8	2.7	88421	9	US-09-976-059-1	Sequence 1, Appl1
36	41.4	2.7	2269	10	US-09-860-298-1	Sequence 1, Appl1
37	41.4	2.7	8280	10	US-09-764-878-384	Sequence 384, App
38	41.4	2.7	1835	10	US-09-764-878-385	Sequence 385, App
39	41.2	2.7	492	9	US-09-854-133-323	Sequence 323, App
40	41.2	2.7	492	10	US-09-738-973-323	Sequence 34, Appl
41	41.2	2.7	1079	10	US-09-804-682-34	Sequence 3, Appl1
42	40.8	2.6	2351	10	US-09-803-589-3	Sequence 724, App
43	40.8	2.6	2500	10	US-09-954-456-724	Sequence 1169, Ap
44	40.8	2.6	2500	10	US-09-954-456-1169	Sequence 1827, Ap
45	40.8	2.6	2500	10	US-09-954-456-1827	

ALIGNMENTS

RESULT 1
US-10-044-090-647
; Sequence 647, Application US/10044090
; Patient No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044, 090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 647
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1519595CB1
US-10-044-090-647

Query Match 3.3%; Score 52; DB 12; Length 2782;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1502 CGCCCGCTCCCGCTCGCTAGCTGCTGCTGCGCGTCCCGCTCC 1553
DB 1 CGCCCGCTCCCGCTCGCTAGCTGCTGCTGCGCGTCCCGCTCC 52

RESULT 2
US-10-025-380-175
; Sequence 175, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong

```

APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skelky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darlick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 175
LENGTH: 407
TYPE: DNA
ORGANISM: Homo sapiens
US-10-025-380-175

```

Query Match	3.1%	Score 47.6;	DB 9;	Length 407;
Best Local Similarity	92.6%	Pred. No. 0.0014;		
Matches 50; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

DY 1500 GCCGCCCGCTTGCCGTCGGTGCCTAGCTGCCTCGGTGCGGCTGTCGCCGCTCC 1553
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GGCACGACTTGCCGTCGGTGCCTAGCTGCCTCGGTGCGGCTGTCGCCGCTCC 54

```

RESULT 3
US-09-922-217-175
: Sequence 175, Application US/09922217
: Patent No. US20020076414A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Lodes, Michael J.
: APPLICANT: Secretist, Heather
: APPLICANT: Benson, Darin R.
: APPLICANT: Meagher, Madeleine Joy
: APPLICANT: Stolk, John A.
: APPLICANT: Wang, Tongtong
: APPLICANT: Jiang, Yugu
: APPLICANT: Smith, Carole Lynn
: APPLICANT: King, Gordon E.
: APPLICANT: Wang, Aijun
: APPLICANT: Clapper, Jonathan D.
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
: TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
: FILE REFERENCE: 210121.471C13
: CURRENT APPLICATION NUMBER: US/09/922,217
: CURRENT FILING DATE: 2001-08-03
: NUMBER OF SEQ ID NOS: 1124
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 175
: LENGTH: 407
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-922-217-175

```

```

Query Match      3.1%; Score 47.6; DB 10; Length 407;
Best Local Similarity 93.6%; Pred. No. 0.0014;
Matches 50; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1500 GCGCCGCGCTGCCGCTGCGTAGCTGCTGCTGGTGGCGCTGCTCCGCTCC 1553
bb      1 GGCACGACCTTGCCTGCTGCTGCTAGCTGCTGCTGGTGGCTGCTCCGCTCC 54

```

RESULT 4
US-09-833-263-175
; Sequence 175, Application US/09833263

```

Patent No. US20020110547A1
GENERAL INFORMATION:
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Stolk, John A.
APPLICANT: Meagher, Madeleine J.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121,471C12
CURRENT APPLICATION NUMBER: US/09/833,263
CURRENT FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 1093
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 175
LENGTH: 407
TYPE: DNA
ORGANISM: Homo sapien
US-09-833-263-175

```

Query Match	3.1%;	Score 47.6;	DB 10;	Length 407;
Best Local Similarity	92.68;	Pred. No. 0.0014;		
Matches 50;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

QY 1500 GCCGCCCGCTTGCCGTCGGTGCCTAGCTGCCTCGGTGCGCCTGCCGCTCC 1553
| | | |||||
Db 1 GGCACGACTTGCCTCGGTGCCTAGCTGCCTCGGTGCGCCTGTTCCCCGCTCC 54

```

RESULT 5
US-10-025-380-121
Sequence 121, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aljun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skelky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 121
LENGTH: 2691
TYPE: DNA
ORGANISM: Homo sapiens
US-10-025-380-121

```

Query Match:	3.0%;	Score 47;	DB 9;	Length 2691;
Best Local Similarity:	100.0%;	Pred. No. 0.0057;		
Matches 47;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1507	GCCTTGGCCGCGGCTCCCTAGCTGCGTGGTGGCGGCTGTCCTCCCGCTCC	1553
Db	1	GCCTGCCCCGCGTCCGCTAGCTGCGTGGTGGCGGCTGTCCTCCCGCTCC	47

RESULT 6
US-09-922-217-121

DB 39270 CCCCCGGGCGCCAGCTCCGGAGCTCAGCGGAGCCAGCCAGCGGCTCCGCGAGACC 39329
QY 1540 CGTGTCTCCGC 1550
DB 39330 CCCTTCCCGCG 39340

RESULT 10

US-09-904-968A-1
; Sequence 1, Application US/09904968A
; Publication No. US20030008288A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: GERMINO, Gregory
; APPLICANT: WATNICK, Terry
; APPLICANT: PHAKDEKITCHAROEN, Bunyong
; TITLE OF INVENTION: DETECTION AND TREATMENT OF POLYCYSTIC KIDNEY DISEASE
; FILE REFERENCE: JHU1680-2
; CURRENT APPLICATION NUMBER: US/09/904,968A
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/283,691
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/218,261
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 53522
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-968A-1

Query Match 3.0%; Score 46; DB 9; Length 53522;
Best Local Similarity 49.7%; Pred. No. 0.054;
Matches 144; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

QY 126 CTCCTTTCCTCCCTTCCTCCCTGATACAGACTCCAAATTCATGAGACTATTAACCTCT 185
DB 34584 CTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34643
QY 186 ACACCCCTCAGATTGCGCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 245
DB 34644 TCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34703
QY 246 TCCCATGATGCTCCCTCCCTGATGCTCCCTGATGCTCCCTGATGCTCCCTGATGCT 305
DB 34704 CCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34763
QY 306 GCGATGTCACCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 364
DB 34764 CCCATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34823
QY 365 CTACCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 414
DB 34824 CTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34873

RESULT 11

US-09-944-413-6
; Sequence 6, Application US/09944413
; Patent No. US20020156004A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin

APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,413
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: NO. US20020156004A1
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: NO. US20020156004A1
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000

LENGTH: 3441
TYPE: DNA
ORGANISM: Homo Sapien
US-09-944-403-6

Query Match
Best Local Similarity 51.0%; Score 42.8; DB 9; Length 3441;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1325 GTGCGGAGCGGCTGTGCTCCAGGCGCAGGCGGCTGTGCGAGCGCCAGGC 1384
DB 8 GTGGGCGGAGCGGTGGGCGCCGCGCCAGCCCGCCGCGCCCTCCGCTCCGCACTC 67
QY 1385 GGGCGCGCTCTTTCAGGCGCGCCAGCTCCCGCCCTGCGCTCCCTCC 1444
DB 68 GCGCTCTCCCTCCCTCCCGCGCTCCCGCTCCCTCCCTCCCTCCAGCTGTC 127
QY 1445 GCTCGAGCTTACTTAACCTGCGCGCGCGCGCGCTCTCACTTCCCTGAGCCG 1504
DB 128 GTTCGGGTATGCGAGAGCTCCCGCGCGCGCGCGCGCTGCTCTGCGGCTGCTG 187
QY 1505 CCGCTTGCGCGCTGCGTGC 1522
DB 188 CTGCTGCGCTCCGCGCG 205

RESULT 13
US-09-944-896-6
Sequence 6, Application US/09944896
Patent No. US20020168715A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017

PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 6
LENGTH: 3441
TYPE: DNA
ORGANISM: Homo Sapien
US-09-944-896-6

Query Match
Best Local Similarity 51.0%; Score 42.8; DB 9; Length 3441;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1325 GTGCGGAGCGGCTGTGCTCCAGGCGCAGGCGGCTGTGCGAGCGCCAGGC 1384
DB 8 GTGGGCGGAGCGGTGGGCGCCGCGCCAGCCCGCCGCGCCCTCCGCTCCGCACTC 67
QY 1385 GGGCGCGCTCTTTCAGGCGCGCCAGCTCCCGCCCTGCGCTCCCTCC 1444
DB 68 GCGCTCTCCCTCCCTCCCGCGCTCCCGCTCCCTCCCTCCAGCTGTC 127
QY 1445 GCTCGAGCTTACTTAACCTGCGCGCGCGCGCGCTCTCACTTCCCTGAGCCG 1504

Accession	Sequence	Position
Dd	128 GTTCCGCGTCATGCGCCGAGCGTCCCGGGCCCCCGGGGGCCCCCGCTGCTGCTCTCTCGGGCTGCTG	187
Qy	1505 CCGCTTTGCCCGCTCGAGTCG	1522
Dd	188 CTGCTCGGCTCCCGGGCCG	205

```

RESULT 14
US-09-944-944-6
: Sequence 6, Application US/09944944
: Patent No. US20020173463A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin
: APPLICANT: Botstein, David
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Flivaroff, Ellen
: APPLICANT: Gerritsen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul
: APPLICANT: Grimaldi, Christopher
: APPLICANT: Gurney, Austin
: APPLICANT: Hillan, Kenneth
: APPLICANT: Kljavin, Ivar
: APPLICANT: Napier, Mary
: APPLICANT: Roy, Margaret
: APPLICANT: Tumas, Daniel
: APPLICANT: Wood, William
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P2548P1C1
: CURRENT FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 09/866,028
: PRIOR FILING DATE: 2001-05-25
: PRIOR APPLICATION NUMBER: 60/067,411
: PRIOR FILING DATE: December 3, 1997
: PRIOR APPLICATION NUMBER: 60/069,334
: PRIOR FILING DATE: December 11, 1997
: PRIOR APPLICATION NUMBER: 60/069335
: PRIOR FILING DATE: December 11, 1997
: PRIOR APPLICATION NUMBER: 60/069,278
: PRIOR FILING DATE: December 11, 1997
: PRIOR APPLICATION NUMBER: 60/069,425
: PRIOR FILING DATE: December 12, 1997
: PRIOR APPLICATION NUMBER: 60/069,696
: PRIOR FILING DATE: December 16, 1997
: PRIOR APPLICATION NUMBER: 60/069,694
: PRIOR FILING DATE: December 16, 1997
: PRIOR APPLICATION NUMBER: 60/069,702
: PRIOR FILING DATE: December 16, 1997
: PRIOR APPLICATION NUMBER: 60/069,870
: PRIOR FILING DATE: December 17, 1997
: PRIOR APPLICATION NUMBER: 60/069,873
: PRIOR FILING DATE: December 17, 1997
: PRIOR APPLICATION NUMBER: 60/068,017
: PRIOR FILING DATE: December 18, 1997
: PRIOR APPLICATION NUMBER: 60/070,440
: PRIOR FILING DATE: January 5, 1998
: PRIOR APPLICATION NUMBER: 60/074,086
: PRIOR FILING DATE: February 9, 1998
: PRIOR APPLICATION NUMBER: 60/074,092
: PRIOR FILING DATE: February 9, 1998
: PRIOR APPLICATION NUMBER: 60/075,945
: PRIOR FILING DATE: February 25, 1998
: PRIOR APPLICATION NUMBER: 60/112,850
: PRIOR FILING DATE: December 16, 1998
: PRIOR APPLICATION NUMBER: 60/113,296
: PRIOR FILING DATE: December 22, 1998
: PRIOR APPLICATION NUMBER: 60/146,222
: PRIOR FILING DATE: July 28, 1999
: PRIOR APPLICATION NUMBER: PCT/US98/19330

```

```

      PRIOR FILING DATE: September 16, 1998
      PRIOR APPLICATION NUMBER: PCT/US98/225108
      PRIOR FILING DATE: December 1, 1998
      PRIOR APPLICATION NUMBER: 09/216,021
      PRIOR FILING DATE: December 16, 1998
      PRIOR APPLICATION NUMBER: 09/218,517
      PRIOR FILING DATE: December 22, 1998
      PRIOR APPLICATION NUMBER: 09/254,311
      PRIOR FILING DATE: March 3, 1999
      PRIOR APPLICATION NUMBER: PCT/US99/12252
      PRIOR FILING DATE: June 22, 1999
      PRIOR APPLICATION NUMBER: PCT/US99/21090
      PRIOR FILING DATE: September 15, 1999
      PRIOR APPLICATION NUMBER: PCT/US99/28409
      PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
      PRIOR APPLICATION NUMBER: PCT/US99/28313
      PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
      PRIOR APPLICATION NUMBER: PCT/US99/28301
      PRIOR FILING DATE: December 1, 1999
      PRIOR APPLICATION NUMBER: PCT/US99/30095
      PRIOR FILING DATE: December 16, 1999
      PRIOR APPLICATION NUMBER: PCT/US00/03565
      PRIOR FILING DATE: February 11, 2000
      PRIOR APPLICATION NUMBER: PCT/US00/04414
      PRIOR FILING DATE: February 22, 2000
      PRIOR APPLICATION NUMBER: PCT/US00/05841
      PRIOR FILING DATE: March 2, 2000
      PRIOR APPLICATION NUMBER: PCT/US00/08439
      PRIOR FILING DATE: March 30, 2000
      PRIOR APPLICATION NUMBER: PCT/US00/14042
      PRIOR FILING DATE: May 22, 2000
      PRIOR APPLICATION NUMBER: PCT/US00/20710
      PRIOR FILING DATE: July 28, 2000
      PRIOR APPLICATION NUMBER: PCT/US00/32678
      PRIOR FILING DATE: December 1, 2000
      PRIOR APPLICATION NUMBER: PCT/US01/06520
      PRIOR FILING DATE: February 26, 2001
      NUMBER OF SEQ ID NOS: 120
      SEQ ID NO 6
      LENGTH: 3441
      TYPE: DNA
      ORGANISM: Homo Sapien
      US-09-944-944-6

      Query Match      2.88;      Score 42.8;      DB 9;      Length 3441;
      Best Local Similarity 51.0%;      Pred. No. 0.096;
      Matches 101;      Conservative 0;      Mismatches 97;      Indels 0;      Gaps

      QY      1325      GTGGCGGAGCGGGGCGGTGGCGCCACAGGCGAGCGGAGGCGGACGGGTTGGAGGCGCCAGGC      1381
      DB      8      GTGGCGGAGCGCGTGGGCGCCGSCACACCGCCCGCGCGGCGCCCTCGGCGTCCGAGACTT      67

      QY      1385      GACCGCGCCCTCTTTCGACAGGGGCGGCGACACTTCCCGCCCTGGGCGTCCGCTCCCTCC      1444
      DB      68      GCGCGCTCCCTCCCTCCCGCGCGGCTCCCGCGGCGCTCCCTCCCTCCCTCCCGACACTCTCC      127

      QY      1445      GCTGCGACACTTAACTTAACTTGGCGCGGCGGAGGCGGAGGCGCTCACTTCCCTGAGAGCGC      1504
      DB      128      GTTGGCGCTGATGCGAGACTTCCGAGCGCCCGCGCCCGCGCTGCTGCTCTCCGCGGCTGCTG      187

      QY      1505      CCGCTTCCCGCTCGGCTCG      1522
      DB      188      CTGCTCGGCTCCGCGCGC      205
  
```

RESULT 15
US-09-944-907-6
Sequence 6, Application US/09944907
Publication No. US20020198147A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan

